Structural Characterization of a Putative GTP-binding Protein, EngB

CHAN, Kwok Ho

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Philosophy
in
Molecular Biotechnology

November 2007


香港中文大學
圖書館系統
13 JAN 2009
UNIVERSITY
LIBRARY SYSTEM
THE CHINESE UNIVERSITY OF HONG KONG

Thesis committee:

Professor Shaw Pang-Chui (Committee chairman)

Professor Wong Kam-Bo (Thesis supervisor)

Professor Tsui Kwok-Wing, Stephen (Committee member)

Professor Zhu Guang (External examiner)

Statement

All experiment works reported here were conducted by the author, unless stated the otherwise.

CHAN Kwok-Ho

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. K.B. Wong, for his guidance and opportunities. Under his supervision, I have learnt a lot from him, from experimental techniques and knowledge to logical thinking and problem solving skills.

Here, I also express my thanks to all my colleagues in MMW507. Your advices and discussion help me a lot through the years here.

Abstract

EngB belongs to a family of novel GTP-binding protein that is essential to the survival of bacteria, and thus, is potential drug targets for the development of antibiotics. However, the function of EngB is poorly understood. Biochemical study showed that EngB family member, YihA in *E. coli*, was capable to bind to guanine nucleotide. Morevoer, EngB in *Escherichia coli* and *Bacillus subtilis* were found to be essential in corresponding species in knock-out experiments. EngB from *Thermotoga maritima* was studied in this work. Crystal of EngB from *T. maritima* has been obtained and its structure was solved. Crystal structures of EngB in complex with GDP and its apo form were obtained. Its structure was compared with the already solved homolog from *B. subtilis*. The switch I region of EngB from *T. maritima* was detected in the crystal structure which was not detected in the *B. subtilis* structure. Thermostability of *T. maritima* EngB was examined. By comparing the crystal structures of EngB from thermophilic *T. maritima* with mesophilic *E. coli* and *B. subtilis*, unique charge-charge interaction and proline residues in *T. maritima* EngB were identified which may bring to further study of thermostability of *T. maritima* EngB. To further characterize the protein EngB, interacting partners of *E. coli* EngB has been isolated by tandem affinity pull-down. Identification of the isolated proteins may help us to understand the biological role of EngB.

摘要

EngB 屬於一個從序列擬測的鳥苷酸結合蛋白 (簡稱：G 蛋白) 家族，我們已知它對細菌的生存的必要性，因此它是一個潛在的抗生素藥物研發對象。不過，對於 EngB 的實際功能，我們所知不多。現在已知的包括，大腸桿菌的 EngB 確實能夠跟鳥苷二磷酸 / 三磷酸結合，另外在大腸桿菌和枯草芽胞桿菌中的基因敲除實驗中，EngB 對相關的生物的生存是必要的。這個研究的對象是海棲熱袍菌的 EngB 蛋白。海棲熱袍菌 EngB 蛋白被表述，純化，然後拿去作結晶，從而解構了海棲熱袍菌 EngB 的元白蛋和鳥苷二磷酸 / EngB 白蛋的複合晶體結構。從海棲熱袍菌 EngB 的鳥苷二磷酸複合晶體結構中，我們解構了從枯草芽胞桿菌 EngB 裡無法解析的 G 蛋白開關一區域。另外，我們還比較了嗜熱的海棲熱袍菌，跟屬於中溫菌的大腸桿菌和枯草芽胞桿菌的 EngB 的結構，發現了在嗜熱的海棲熱袍菌 EngB 獨有的幾組帶電殘基之間的電荷作用和脯胺酸分佈，它們都可能對海棲熱袍菌 EngB 的嗜熱性有增強作用。另外，爲了深化對 EngB 功能的認識，發展了一個並列親合蛋白質親和力試驗，用來提純跟大腸桿菌 EngB 有相互作用的蛋白，識別到這些蛋白對認識 EngB 的功能有重大幫助。

Table of Contents

Chapter 2 Materials and methods

Chapter 4 Structural details of TM EngB

Chapter 5 Purified TM EngB is Active for binding guanine nucleotide but inactive for GTPase hydrolysis activity

Abbreviations of amino acids

A	Ala	Alanine
B	Asx	Asparagine or aspartate
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
X	-	Unknown or any amino acid
Y	Tyr	Tyrosine
Z	Glx	Glutamine or glutamate

Abbreviations of species name

TM	*Thermotoga maritima*
BS	*Bacillus subtilis*
EC	*Escherichia coli*

Chapter One

1 General Introduction

This piece of work focuses on the study of a putative protein EngB from a hyperthermophilic bacteria *Thermotoga maritima*. This work aims to enrich our understanding towards EngB. Crystal structure of *T. maritima* EngB with the already solved structure of *B. subtilis* YsxC will provide information about the conservation of structure-function of the protein family. The experimental characterization of these conserved hypothetical proteins is expected to reveal new, crucial aspects of microbial biology and lead to identification of novel anti-microbial drug targets.

1.1 GTPase in general

GTPases are a large family of hydrolase enzymes that can bind and hydrolyze GTP. The GTP binding and hydrolysis takes place in the highly conserved G domain common to all GTPases. The hydrolysis of the γ phosphate of GTP into guanosine diphosphate (GDP) and orthophosphate occurs by the S_N2 mechanism via a pentavalent intermediate state and is dependent on the magnesium ion Mg^{2+}. GTPases play an important role in signal transduction at the intracellular domain of transmembrane receptors, protein biosynthesis at the ribosome, control and differentiation during cell division, translocation of proteins through membranes and transport of vesicles within the cell.

1.2 G proteins and GTP switch

G proteins are regulatory GTPases. G proteins, short for guanine nucleotide binding proteins, are a family of proteins involved in second messenger cascades. They are so called because of their signaling mechanism, which uses the exchange of guanosine diphosphate (GDP) for guanosine triphosphate (GTP) as a general molecular "switch" function to regulate cell processes.

Throwing the switch is performed by the unidirectional change of the GTPase from the active, GTP-bound form to the inactive, GDP-bound form by hydrolysis of the GTP through intrinsic GTPase-activity, effectively switching the GTPase off. During the process, conformation of the switch region of the protein changed, and such change may alter the function of other proteins and thus a signal is passed from one protein to another through the altering of switch state to initiate other cellular process.

The altering of switch states may require presence of other proteins, they are the GTPase-activating proteins (GAPs) and guanine nucleotide exchange factors (GEFs). The hydrolysis reaction of on-state GTP-bound G protein is initiated by GAPs, coming from upstream signal transduction pathway. The GDP-bound G protein can be reverted (switching the GTPase on again) by GEFs, which cause the GDP to dissociate from the GTPase, leading to its association with a new GTP. This closes the cycle to the active state

of the GTPase; the irreversible hydrolysis of the GTP to GDP forces the cycle to run only in one direction. Therefore only the active state of the GTPase can transduce a signal to a reaction chain. Interfering the interaction between the G proteins with their corresponding GAPs and GEFs may lock the related pathway in one state.

1.3 Structural similarities in GTPases

GTPases share a set of conserved G box GDP/GTP-binding motif elements beginning at the N-terminus: G1, GXXXXGKS/T; G2, T; G3, DXXGQ/H/T; G4, T/NKXD; and G5, C/SAK/L/T [Bourne, H.R. *et al.*, 1991]. Besides, two surface loops, switch I and switch II are present and they undergo conformational changes upon binding of GTP or GDP nucleotides. Such change of switch I and switch II are responsible for switching a biochemical process from one state to another. To pass the signal of switching a biochemical process, the GTPase switch will physically contact the downstream proteins. Loops conformation serve as a switch to turn "on" or "off" the downstream protein which receive the signal given by the switch state. Together, these elements make up an ~20 kDa G domain that has a conserved structure and biochemistry shared by all GTPases.

1.4 G proteins in bacteria

Bacteria do not use G proteins to regulate membrane signaling pathways like

eukaryotes. There is no analogue of the α-subunit of heteromeric G proteins which is essential in transmitting signals from membrane receptors. Also, the small GTPase subfamilies (Ras, Rho, ARF, Rab and Ran) seem to be absent. It may be that prokaryotes accomplish membrane signaling largely through two-component regulators [Caldon *et al.*, 2001]. Current knowledge suggest that bacterial G proteins have functions linked to the ribosome and nucleic acid binding [Caldon and March, 2003]. Other than that, bacteria usually contain far less molecular switch G proteins. Bacteria with large genomes contain 20–30 molecular switch GTPases, whereas bacteria with small genomes contain as few as 11 [Fraser *et al.*, 1995].

1.5 Background information of the protein family EngB

With arising number of completed genome sequencing projects, more and more genes are found with unknown function. Many of them across different species can be clustered into groups or families of proteins according to their sequence similarities and presence of particular domain. Many of them are regarded as hypothetical proteins as they are predicted to be translated and transcribed with open-reading frame. Among these proteins, a number of putative GTPases, some of them control essential function in cells, have been identified [Caldon, C.E., *et al.*, 2003]. EngB family is one of the hypothetical GTPase families. This family is named as YihA family, after the *E. coli* member, or EngB,

essential neisserial GTPase. By amino acid sequence analysis, they have been classified as member in the translation factor related (TRAFAC) class of GTPase superfamily [Lipe, D.D., *et al.*, 2002]. The TRAFAC GTPases are characterized by a highly conserved threonine or serine residue in the loop between strands 2 and 3, that makes a hydrogen bond to the Mg^{2+} cation required for GTP hydrolysis.

Orthologues of EngB are found in all three domains of life. *In vitro* study of family member, YihA in *E.coli*, suggested that this group of protein can actually bind to guanine nucleotide [Lehoux *et al.*, 2003]. In parallel to our work, crystal structures of the mesophilic YsxC [Ruzheinikov *et al.*, 2004], YihA and EngB from *Pyrococcus Horikoshii* were already solved. What makes EngB more interesting is this gene is essential for growth in *E.coli* and *B. subtilis*. YihA and YsxC were found to be essential in corresponding species in knock-out experiments [Arigoni *et al.*, 1998; Dassain *et al.*, 1999; Pragai and Harwood, 2000]. In the knock-out experiments of YihA in *E. coli*, cell cycle defect was observed [Dassain *et al.*, 1999]. Therefore, it is speculated that the EngB family is related to cell cycle process. The broad-spectrum nature of EngB and its essentiality for cell viability in bacteria make it an attractive antibacterial target.

1.6 Basic information of EngB in *Thermotoga maritima*

No previous study has been performed on EngB in *T. maritima* (TM EngB). TM EngB is a basic protein with theoretical pI of 9.7 and has a molecular weight of 22kDa. It consists of 194 amino acids.. EngB from *T. maritima* is chosen because *T. martima* is a hyperthermophilic bacterium. Proteins from thermophlic organism generally behave better due to their tolerance to extreme temperature. Chance for denaturation under normal circumstances is expected to be minimal which help our study by providing quality protein sample in adequate quantity. Besides, there is no structural model of EngB solved from thermophilic source. Structural model of *T. maritima* EngB will be the first EngB from thermophilic source solved and it can be used for further thermostability study with the benefit of small size of the protein (22kDa, <200 residues).

1.7 Objectives of this work

In this work, structure and biochemical properties of *T. maritima* EngB (TM EngB) will be studied. To understand how TM EngB performs its function as a GTP-binding protein, crystal structures of TM EngB will be solved by X-ray crystallography. The mode of binding of guanine nucleotide of TM EngB will be discussed. Structure solved will be compared to the previously solved model to see if extra information can be provided by this new model. The affinity of TM EngB towards its ligand GDP and GTP will be studied by fluorescence titration. Thermostability of TM EngB will be tested by thermal

unfolding and chemical unfolding. Discussion will be made based on the comparison of mesophilic EngB and thermophilic EngB model. It will give us insight on how protein structure is related to the stability of protein at high temperature. In addition, a dual-tag pull-down system which aims at isolating potential interacting partner of EngB in *E. coli* will be established to test the feasibility of finding interacting partner of EngB by pull-down.

Chapter Two

2 Materials and Methods

2.1 Materials

2.1.1 Chemical Reagents

Table 2.1 List of common chemicals

Chemical Reagents	Formula Weight	Source
2-mercaptoethanol (β-ME) (C_2H_6OS)	78.1	Sigma, USA
Acrylamide / Bis-acrylamide; 30% solution, mix solution 37.5:1	/	Sigma, USA
Agarose powder	/	Sigma, USA
Ammonium Persulfate (APS) ($(NH_4)_2S_2O_8$)	228.2	Serva, USA
Ampicillin, Sodium Salt	371.4	USB, USA
Bromophenol Blue ($C_{19}H_9Br_4O_5SNa$)	691.9	Sigma, USA
Chloramphenicol ($C_{11}H_{12}Cl_2N_2O_5$)	323.13	Sigma, USA
Coomassie Brilliant Blue ($C_{45}H_{44}N_5O_7S_2Na$)	826.0	USB, USA
Deoxynucloetides (dNTP)	/	Novagen, Germany

Ethanol (CH_3CH_2OH)	46.1	Merck, Germany
Ethidium Bromide (EtBr)	394.0	Amresco, USA
Ethylenediaminetetraacetic Acid (EDTA) ($C_{10}H_{16}N_2O_8$)	292.2	Sigma, USA
Guanidine hydrochloride (CH_6N_3Cl)	95.53	Sigma, USA
Glycine ($H_2NCH_2CO_2H$)	75.1	USB, USA
Hydrochloric Acid (HCl)	36.5	BDH, England
Isopropanol ($(CH_3)_2CHOH$)	60.1	BDH, England
Luria broth	/	USB, USA
Methanol (CH_3OH)	32.1	Merck, Germany
N,N,N',N-Tetramethyl-Ethyl ene - Diamine (TEMED) ($(CH_3)_2NCH_2CH_2N(CH_3)_2$)	116.2	Bio-Rad, USA
Phenylmethyl-sulfonyl Fluoride (PMSF) ($C_7H_7FO_2S$)	174.2	Sigma, USA

Post-stained SDS-PAGE Standards (low range)	/	Bio-Rad, USA
Potassium Chloride (KCl)	74.6	USB, USA
Potassium Phosphate (KH_2PO_4)	136.1	Sigma, USA
Sodium acetate (NaAc) (CH_3COONa)	82.03	Sigma, USA
Sodium Chloride (NaCl)	58.4	USB, USA
Sodium Dodecyl Sulfate (SDS) ($CH_3(CH_2)_{11}SO_4Na$)	288.4	USB, USA
Sodium Hydrogen Phosphate (Na_2HPO_4)	142.0	Sigma, USA
Sodium Hydroxide (NaOH)	40.0	Sigma, USA
Tris ($NH_2C(CH_2OH)_3$)	121.1	USB, USA

2.1.2 Buffers

2.1.2.1 Preparation of Buffers

Buffer was prepared by dissolving chemicals in distilled water (dH_2O) and adjusted to suitable pH using either HCl or NaOH, unless otherwise specified.

2.1.2.2 Buffers for Common Use

RF1 (competent cell preparation)

30mM KAc, 100mM $RbCl_2$, 10mM $CaCl_2$, 50mM $MnCl_2$, 15% glycerol. pH was adjusted to 5.8 by HAc and no back titration was performed. The solution was sterilized by filtration through a 0.2μm filter.

RF2 (competent cell preparation)

10mM MOPS, 75mM $CaCl_2$, 10mM $RbCl_2$, 15% glycerol. pH was adjusted to 6.5 with KOH and the solution was sterilized by filtration through a 0.2μm filter.

6X DNA loading buffer

0.25% (w/v) bromophenol blue, 40% (w/v) sucrose and 0.25% (w/v) xylene cyanol FF. The buffer was stored at 4°C.

50X TAE buffer stock

2 M Tris-acetate and 50 mM EDTA

SDS running buffer

25 mM Tris base, 192 mM glycine and 0.1% (w/v) SDS.

6X SDS loading buffer

12% (w/v) SDS, 375 mM Tris-HCl, 60% (v/v) glycerol, 0.01% bromophenol blue, 5% (v/v) β-ME. pH was adjusted to 6.8. and the buffer was stored at 4°C.

Coomassie brilliant blue staining solution

Acetic acid, ethanol and dH_2O were mixed in a ratio of 1:3:10. 0.05% (w/v) Coomassie brilliant blue R250 was then added.

Destain solution was prepared with 40% (v/v) ethanol and 10% (v/v) acetic acid.

LB broth

25g LB broth powder (USB) was added to 1L dH_2O. The solution was sterilized by autoclaved.

100μg/ml ampicillin was added for LBA medium,

50μg/ml chloramphenicol was added for LBC medium,

100μg/ml ampicillin and 50μg/ml chloramphenicol were added for LBAC medium.

LB agar plate

25g/L LB powder (USB) and 20g/L agar powder (USB) was added to 1L dH_2O. The solution was sterilized by autoclaved. After the agar was cooled to hand-hot, corresponding antibiotics were added and 20mL LB agar was poured into each 100mm diameter plates.

100μg/ml ampicillin was added for LBA agar plate

50μg/ml chloramphenicol was added for LBC agar plate

100μg/ml ampicillin and 50μg/ml chloramphenicol were added for LBAC agar plate

40% glucose (dextran)

400 g/L glucose was dissolved in hot dH_2O gradually with continuous stirring. The solution was sterilized by filtering through 0.2uM filter (Millipore).

Buffer for circular dichroism experiment

20mM phosphate pH 7.5. Prepared buffer was filtered by 0.2uM filter (Millipore)

20mM phosphate pH7.5, 8M guanidine hydrochloride (GdnHCl). The buffer concentration was determined by refractive index. $[GdnHCl]= 57.147*(N_{sample} - N_{buffer}) + 38.68*(N_{sample} - N_{buffer})^2 - 91.6*(N_{sample} - N_{buffer})^3$ [Nozaki, Y., 1972]

Prepared buffer was filtered by 0.2uM filter (Millipore)

2.1.3 Expression Strains and Plasmids

Escherichia coli (*E. coli*) strain DH5α (Novagen) was used for plasmid DNA amplification.

E. coli strain C41 (DE3) (Novagen) was used for protein expression.

2.1.4 Primer list

The following listed the sequences of the forward (F) and reverse (R) primers from 5' to 3':

TM EngB (F)

TATGCAGGATCCATTATCAGAGATGTAGAA

TM EngB (R)

TATGCAGAATTCTCAATTTTCTTTCAGTAA

SBP-CBP-NdeI (F)

TATGCACATATGGACGAGAAGACCACC

SBP-CBP-LVPRGS-BamHI (R)

TATGCAGGATCCACGCGGAACCAGAAGTGCCCCGGAGGATGAG

Ecoli EngB (F)

TATGCACATATGTTGACTAATTTGAATTATC

Ecoli EngB (R)

TATGCAGAATTCTTATTCGCCGTCCTGCGTTTC

2.2 Methods

2.2.1 Preparation of competent cells

Competent cells *E.coli* DH5α or BL21 pLysS (DE3) from -80°C glycerol stock were streaked to an LB agar plate (DH5α) or LBC agar plate (BL21 pLysS) and the plate was incubated at 37°C overnight. Starter culture was prepared by picking a single colony and inoculated into 5ml LB medium or LBC medium. The culture was shaken at 37°C with 280 rpm until OD600 reached 0.3. Then, the culture was transferred into 100ml medium and continued shaking at 37°C until OD600 reached 0.4-0.6. The cells were placed on ice for 5 minutes. The bacteria were collected by centrifugation at 8000 rpm for 10 min at 4°C (Beckman JA-16.250 rotor). The cell pellet was suspended in 40ml RF1 and kept on ice for 5 minutes. Afterward, the cells were re-centrifuged and the pellet was suspended in 4ml RF2. The suspension was kept on ice for 15 min and then divided into 100μl aliquots.

2.2.2 Cloning

2.2.2.1 Cloning of target genes by PCR

The target gene was cloned from genomic DNA (ATCC) by PCR. A pair of primers was designed to anneal to the DNA template. The primers acted in a forward (5' to 3' with respect to the reading frame) and reverse (3' to 5' with respect to the reading frame) pair.

PCR was performed to amplify the target gene. Composition of reaction mixture and the general reaction scheme are listed in Tables 2.2 and 2.3 respectively. The reactions were carried out by pfu DNA polymerase (Promega) in 0.2mL PCR tubes on thermal cycler. The resulting DNA products were analyzed by agarose gel electrophoresis or stored at -20°C.

Table 2.2 PCR reaction mixture (1X)

Distilled water	39 µl
10X pfu buffer	5 µl
10mM dNTPs (2.5 mM each of dATP, dCTP, dGTP and dTTP)	2 µl
Forward primer (25 µM)	1 µl
Reverse primer (25 µM)	1 µl
DNA Template	1 µl
pfu DNA polymerase (5 U/µl)	1 µl
Total	50 µl

Table 2.3 PCR cycles

		Temperature	Time	Number of cycles
Initial denaturation		95°c	5 min	---
Thermal cyclel	Denaturation	95°c	30 sec	35
	Annealing	50°c	30 sec	
	Elongation	72°c	2 min	
Final extension		72°c	10 min	---
storage		4°c	∞	---

2.2.2.2 Agarose Gel Electrophoresis

1.0 % (w/v) agarose gel was prepared by dissolving 0.4 g of agarose in 40 ml 1X TAE buffer, with 2 µl of ethidium bromide (10 mg/ml) added. PCR products were mixed with 6X loading buffer to give a final concentration of 1X. The gel was submerged into a gel tank containing 1X TAE buffer. After loading the samples and DNA Ladder Mix into the wells, electrophoresis was performed with a constant voltage of 100 V. The agarose gel was then placed onto a transilluminator (300 nm) for visualization of DNA.

2.2.2.3 Extraction and Purification of DNA from Agarose Gels

The protocol was modified from the manual provided by the manufacturer (Gel-M Kit, Viogene). Desired DNA fragment was excised from the gel under the illumination of a transilluminator. 500 µl of GEX buffer was added to the centrifuge tube containing the gel slice. The mixture was incubated in 70°C until the gel completely dissolved. A column was placed onto a collection tube and the gel mixture was loaded onto the column. The column was then centrifuged at 13,000 rpm by desktop centrifuge for 60 s and the flow-through was discarded. Next, the column was washed with 500 µl of WF buffer and then centrifuged at 13,000 rpm for 60 s. The flow-through was discarded. After that, the column was washed with 700 µl of WS buffer and then centrifuged at 13,000 rpm for 60 s. The

column was further centrifuged for 3 min to ensure all residual buffer was removed. After that, the column was placed onto a new centrifuge tube and 50 µl of warm water was added onto the column membrane. It was allowed to stand at room temperature for 2 min. DNA was then eluted through centrifugation at 13,000 rpm for 2 min.

2.2.2.4 Restriction Digestion of DNA

For directional cloning, the gene clean product (insert) and the vector (plasmid) were digested by corresponding restriction enzymes (New England Biolabs) in the buffer recommended by the manufacturer at 37°C for 1 hour to overnight. Detailed conditions for restriction digestion were listed in Table 2.4. After digestion, 6X DNA loading dye was added to stop the reaction and the reaction mixture was analyzed by agarose gel electrophoresis as described in section 2.2.2.2. Desired DNA fragment was extracted and purified from the gels by Gel-M kit as described in section 2.2.2.3 for subsequent ligation.

2.2.2.5 Ligation of digested insert and expression vector

The digested PCR products and vector were mixed as shown in Table 2.5 and the reaction mixtures were incubated for ligation at 16°C for at least 16 hours. T4 DNA ligase (New England BioLabs Inc.) was used. The ligation products were transformed into E. coli DH5α competent cells for plasmid amplification.

Table 2.4 Restriction enzyme digestion mixture

	Vector(100ug/mL)	Insert
Undigested vector / insert (PCR product)	10 µl	27 µl
10X OPA buffer (New England BioLabs Inc.)	2 µl	4 µl
Restriction enzyme 1	1 µl	1 µl
Restriction enzyme 2	1 µl	1 µl
dH_2O	6 µl	7 µl
Total	20 µl	40 µl

Table 2.5 Ligation mixture

10x T4 ligase buffer	2 µl
Linearized plasmid vector	5 µl
Digested DNA insert	12 µl
T4 DNA ligase (400 U/µl)	1 µl
10x T4 ligase buffer	2 µl
Total	20 µl

2.2.2.6 Transformation and plating out transformants for Miniprep

10 µl ligation products were incubated with 100µl *E. coli* strain DH5α competent cells for

30 min on ice. A heat shocked at 42°C for 2 minutes and cold shocked for 10 minutes on ice were then applied. 900 µl LB medium was added to the cells and incubated for 1 hour at 37°C with constant shaking at 250 rpm. Afterwards, the mixtures were centrifuged at 13,000 rpm for 1 minute to obtain the bacterial pellet. The pellet was resuspended by 100 µl supernatant and spread on a LBA agar plate. The plate was incubated at 37°C for 16 hours.

2.2.2.7 Verification of insert by PCR

DH5α colonies grown on the plate were picked by toothpicks and subjected to PCR to check if the plasmids had been transformed into the cells. The composition of PCR reaction mixture and the reaction scheme are listed in Table 2.6 and 2.7 respectively. The PCR products were then loaded for agarose gel electrophoresis.

Table 2.6 PCR reaction mixture for verification of insert

10X reaction buffer (New England BioLabs Inc.)	5 µl
$MgCl_2$ (25 mM)	1 µl
dNTP	1 µl
Forward primer (25 µM)	1 µl
Reverse primer (25 µM)	1 µl
Taq polymerase (5 U/ul)	1 µl

dH_2O	40 µl
Total	50 ul

Table 2.7 PCR cycles for clone checking

		Temperature	Time	Number of cycles
Initial denaturation		95°c	5 min	---
Thermal cyclel	Denaturation	95°c	30 sec	35
	Annealing	50°c	30 sec	
	Elongation	72°c	2 min	
Final extension		72°c	10 min	---
Storage		4°c	∞	---

2.2.2.8 Mini-preparation of Plasmid DNA

Single positive clone checked was picked and inoculated into 5 ml of LB medium containing 100 µg/L of ampicillin and incubated at 37°C for 16 hours with constant shaking at 250 rpm. Steps of plasmid extraction were modified from the manual provided by the manufacturer (Wizard Plus SV Minipreps DNA Purification Kit, Promega). First, overnight bacterial culture was centrifuged at 13,000 rpm by desktop centrifuge (Minispin, Eppendorf) for 30 s. Supernatant was discarded. 250 µl of MX1 buffer was added to resuspend the cell pellet. 250 µl of MX2 buffer was then added and the tubes were inverted gently for several times to lyse the cells until the lysate became clear. The mixture was incubated at room temperature for 5 min. After that, 350 µl of MX3 was added. The mixtures were mixed immediately by inversion for several times to neutralize the lysate

until white precipitates formed. The resulting lysate was centrifuged at 13,000 rpm for 10 min. After centrifugation, supernatant was transferred to a DNA-binding column sitting on a collection tube, and centrifuged at 13,000 rpm for 30 s. Flow-through was discarded. After that, 500 µl of WF buffer was added and centrifuged at 13000 rpm for 30 s to wash the column. Flow-through was discarded. 700 µl of WS buffer was then added and centrifuged at 13,000 rpm for 30 s to wash the column. Flow-through was discarded. As any residual ethanol could affect the quality of DNA and inhibit subsequent enzymatic reactions, the column was further centrifuged for 3 min to ensure all residual ethanol had been removed. Finally, the column was transferred to a clean eppendorf tube. 50 µl of warm water was added onto the column membrane. The tube was allowed to stand at room temperature for 2 min, and centrifuged at 13,000 rpm for 2 min. The eluted DNA was stored at -20°C for later use.

2.2.2.9 Confirmation of miniprep product by restriction enzyme digestion

Restriction digestion was performed at 37°C for 2 hours. Details of the reaction mixture are listed in Table 2.8. After digestion, 6X DNA loading dye was added to stop the reaction and the size of insert was checked by agarose gel electrophoresis.

Table 2.8 Restriction digestion mixture for confirmation of miniprep products

dH_2O	14 µl
Miniprep products	4 µl
10X NEB buffer	1 µl
Enzyme 1	0.5 µl
Enzyme 2	0.5 µl
Total	20 µl

2.2.2.10 Sequencing of the plasmid DNA

10µL of plasmid DNA (100µg/mL) was sent to Macrogen Limited to perform sequencing with forward and reverse primers. The sequencing results were aligned with the original target sequence using the program ClustalW [Thompson, J.D., *et al.*, 1994] to check if the sequence contained any undesired mutations.

2.2.3 Expression of the recombinant MBP-TM EngB protein and SBP-CBP EC EngB

2.2.3.1 Transformation for protein expression

Plasmids were transformed to freshly thawed *E. coli* strain BL21(DE3)pLysS competent cells by incubating 1 µl (100µg/mL) plasmids and 100 µl competent cells together on ice for 30 minutes. A heat shock at 42°C for 2 minutes and cold shock for 10 minutes on ice were then applied. The mixtures were spread on LBA plate and incubated at 37°C for 16

hours.

2.2.3.2 Preparation of starter culture

10-20 bacterial colonies was picked and inoculated into 20 ml LB medium with 100 mg/L ampicillin and 50mg/L chroramphenicol. The culture was incubated at 37°C for 3-4 hours with shaking at 280 rpm until OD reached 0.3.

2.2.3.3 Expression of recombinant protein

The starter culture was inoculated into 2L LB culture medium, supplemented with 2ml 1 M $MgSO_4$, 20 ml 40% glucose, 100 mg/L ampicillin and 50mg/L chroramphenicol. The culture medium was then incubated at 37°C at 280 rpm. When OD600 reached 0.4-0.6, 0.4 mM IPTG was added for induction. The broth was allowed to grow for 16 hours at 25°C with shaking at 280 rpm.

2.2.3.4 Cell harvesting

Cells were harvested by centrifugation at 4°C, 8,000 rpm for 10 min (Beckman JA-16.250 rotor). The supernatant was discarded, and the cell pellet was resuspended in sonication buffer or stored at -80°C.

2.2.3.5 Releasing the cell content

All procedures were performed on ice. Cell pellet from 1L of culture was resuspended in 40 ml sonication buffer (20 mM sodium phosphate, 1 mM phenylmethysulfonyl fluoride (PMSF), pH 7.5). The cells were then lysed by sonication for 12 four-seconds bursts, with 1 second intervals and repeated for 5 cycles. The whole lysate was centrifuged at 4°C, 12,000 rpm for 45 min (Beckman JA-25.50 rotor) to remove cell debris. Whole lysate, supernatant and pellet were analyzed by Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE).

2.2.3.6 Check for protein expression by SDS-PAGE

SDS-PAGE was performed using a Mini-PROTEAN II electrophoresis cell (Bio-Rad). Apparatus was set up according to manufacturer's instructions. An inner and an outer glass plates were assembled into a gel cassette using two 0.75 mm thick spacers. 3 ml resolving gel solution was transferred into the space between the inner and outer glass plates. 200 µl of isopropanol was added on top of the gel to ensure a horizontal gel surface and to keep the gel solution away from atmospheric oxygen which inhibits polymerization of acrylamide. After resolving gel polymerization, isopropanol was removed, and stacking gel solution was added on top of the resolving gel. A comb of thickness 0.75 mm with 10 or 15

wells former was inserted into the stacking gel before polymerization. Details of gel compositions are listed in Table 2.9. After polymerization of stacking gel, the comb was removed and the gel cassette was then assembled into an electrophoresis cell. The inner chamber was filled with fresh SDS running buffer; and the outer chamber with re-used SDS running buffer.

6X SDS gel loading buffer was added to the protein samples, followed by denaturation at 95°C for 5 min. Protein samples were then loaded into the wells for electrophoresis. Low molecular weight marker (Amersham) was used as marker. A constant current of 30 mA per gel was applied until the dye front reached the bottom of the gel. After that, the gel was stained with Coomassie brilliant blue staining solution for 1 hour and subsequently destained by destain solution. Finally, the gel was equilibrated in distilled water.

Table 2.9 Composition of resolving and stacking SDS-PAGE gel

Resolving gel	**12.5%**	**15%**
dH_2O	1.16 ml	0.85
30% acrylamide solution (Acrylamide: Bis-acrylamide = 40:1)	1.56 ml	1.875
1.5 M Tris (pH 8.8)	0.95 ml	0.95
10% SDS	38 µl	38µl

10% ammonium persulfate	38 µl	38µl
TEMED	1.5 µl	1.5µl
Total	~ 3.75 ml	~ 3.75 ml

Stacking Gel (5%)	
dH_2O	1.35 ml
30% acrylamide solution (Acrylamide: Bis-acrylamide = 40:1)	0.33 ml
1.5 M Tris (pH 8.8)	0.25 ml
10% SDS	20 µl
APS	20 µl
TEMED	2 µl
Total	~ 2ml

2.2.4 Purification of TM EngB

Purification for TMEngB from soluble cell lysate consists of 4 steps.

A. SP ion exchange chromatography

B. Thrombin digestion to yield non-fusion protein

C. Heparin affinity chromatography

D. Gel filtration (size exclusion) chromatography

The procedures were described as follow

2.2.4.1 SP ion-exchange chromatography

5mL HiTrap™ Sepharose™ HP IEX Columns (Amersham Biosciences) was used in this

chromatography. The SP column was first washed with buffer B (20mM sodium phosphate buffer pH7.5, 1M NaCl) to wash away residual proteins remained in the column. The SP column was equilibrated with buffer A (20mM sodium phosphate buffer pH7.5. The lysate prepared in 2.2.3.5was filtered with a 0.22μm filter and then loaded into the SP column at 4 ml/min. The column was washed with buffer A at 4 ml/min until the OD280 reading at the outlet reached the baseline to remove unbound protein. Proteins were eluted with a linear gradient of 0 to 1 M NaCl in 20 mM sodium phosphate buffer at pH 7.5 over 150ml at 3 ml/min (gradient was made by 0-100% buffer B). Fractions were collected at 6ml per tube. The fractions were analyzed by 12.5% SDS-PAGE to trace for the fractions containing fusion MBP-TM1466

2.2.4.2 Thrombin digestion to remove MBP tag

Thrombin bovine plasma (lyophilized powder 40-300 NIH units/mg protein (biuret)) from Sigma-Aldrich (product code: T4648) was used. The fractions containing the fusion protein after the SP column was pooled together concentrated to make a 20ml solution and filtered with a 0.22μm filter. The absorbance at 280nm (OD280) of the protein solution was measured. For every 1A of OD280, 7mg of the thrombin powder was added to the solution to make the digestion mixture. The digestion mixture was incubated at 37°C overnight. The digestion mixture was analyzed by 12.5% SDS-PAGE to trace for the digestion of fusion MBP-TM EngB to yield MBP and TM EngB.

2.2.4.3 Heparin affinity chromatography

5mL HiTrap™ Heparin™ HP Columns (Amersham Biosciences) was used in this chromatography. Before running a chromatography, the digestion product was subjected to dialysis with buffer A. The solution was injected into a dialysis tubing with molecular weight cut-off= 10,000Da. After sealing the dialysis tubing, it was put into 4L of buffer A at 4°C for 2 hours. It was then filtered with a 0.22μm filter. The Heparin column was washed with buffer B to wash away residual proteins remained in the column. The Heparin column was equilibrated with Buffer A. The filtered supernatant was loaded into the Heparin column at 4 ml/min. The column was washed with buffer A at 4 ml/min until the OD280 at outlet reading reached the baseline to remove unbound protein. Proteins were eluted with a linear gradient of 0.1 to 0.8 M NaCl in 20 mM sodium phosphate buffer at pH 7.5 over 150ml at 3 ml/min (gradient was made 0-80% buffer B). Fractions were collected at 6ml per tube. The fractions were analyzed by 12.5% SDS-PAGE to trace for the fractions containing TM EngB protein

2.2.4.4 Gel filtration chromatography

Superdex 75 (Amersham Biosciences) in XK-26 (60cm x 2.6cm) was used. The fractions containing TM EngB after running heparin column was filtered with a 0.22μm filter. It was then concentrated to about 2 ml using Amicon Ultra-15 Centrifugal Filter Units (Millipore)

with 10,000 MWCO at 3500rpm at 4°C. The column and the 5-ml sample loop were equilibrated with 20mM sodium phosphate, pH7.5, 0.1M NaCl, 1mM EDTA for one bed volume of the gel filtration column. After equilibration, the filtered and concentrated protein was injected through the 5ml sample loop. Elution was done with 20mM sodium phosphate, pH7.5, 0.1M NaCl, 1mM EDTA at 3 ml/min. Fractions were collected at 5ml per tube after the first 90ml elution. The fractions were analyzed by 12.5% SDS-PAGE to find which were the fractions containing TM EngB. The fractions containing TM EngB were pooled and filter with 0.2μm filter and concentrated to OD_{280} ≈10, i.e. 0.33mM (extinction coefficient=29910 for TM EngB) using Amicon Ultra-15 Centrifugal Filter Units (Millipore) with 10,000 MWCO at 3500rpm at 4°C. The TM1466 solution was aliquoted to 200μl per tube and stored at –80°C.

2.2.5 Purification of SBP-CBP-EC EngB

Purification for TMEngB from soluble cell lysate consists of 4 steps.

A. SP ion exchange chromatography

B. Gel filtration (size exclusion) chromatography

The procedures were described as follows

2.2.5.1 SP ion-exchange chromatography

5mL HiTrap™ Sepharose™ HP IEX Columns (Amersham Biosciences) was used in this

chromatography. The SP column was first washed with buffer B (20mM sodium phosphate buffer pH7.5, 1M NaCl) to wash away residual proteins remained in the column. The SP column was equilibrated with buffer A (20mM sodium phosphate buffer pH7.5). The lysate prepared in 2.2.3.5 was filtered with a 0.22μm filter and then loaded into the SP column at 4 ml/min. The column was washed with buffer A at 4 ml/min until the OD280 reading at the outlet reached the baseline to remove unbound protein. Proteins were eluted with a linear gradient of 0 to 0.7 M NaCl in 20 mM sodium phosphate buffer at pH 7.5 over 150ml at 3 ml/min (gradient was made by 0-70% buffer B). Fractions were collected at 6ml per tube. The fractions were analyzed by 12.5% SDS-PAGE to trace for the fractions containing fusion SBP-CBP-EC EngB.

2.2.5.2 Gel filtration chromatography

Procedures for gel filtration chromatography were performed as described in 2.2.4.4

The fractions containing SBP-CBP EC EngB were filtered with 0.2μm filter and concentrated to OD_{280} ≈1, i.e. 30μM (extinction coefficient=33460 for SBP-CBP EC EngB) using Amicon Ultra-15 Centrifugal Filter Units (Millipore) with 10,000 MWCO at 3500rpm at 4°C. The protein solution was aliquoted to 200μl per tube and stored at –80°C.

2.2.6 Protein concentration quantitation

Quantitation of protein was done by measuring the absorbance of protein solution at 280

nm. Extinction coefficient for the protein was estimated by ProtParam (http://au.expasy.org/tools/protparam.html). Concentration of the protein was estimated by applying the Beer's Law **A = εbc**, where,

A = Absorbance

ε = Molar extinction coefficient ($M^{-1}cm^{-1}$)

b = Path length of sample (cm)

c = Concentration of compound in solution (M^{-1})

2.2.7 Crystallography of TM EngB

2.2.7.1 Crystallization preparation

Protein TM EngB was concentrated to ~10mg/mL to prepare crystallization protein stock.

2.2.7.2 Crystallization screening by sitting drop method

Several Hampton Research crystallization kits were screened, including Crystal Screen ™, Crystal Screen 2 ™, Index™. Screening was done by sitting drop experiment. 100 µl of the each crystallization buffer was added to the corresponding reservoir. 1 µl concentrated protein was mixed with equal volume of the a crystallization buffer and added to one of the drop well of the CrystalQuick™ 96 Well Sitting Drop Plate (Hampton research). 1 µl protein buffer 20mM sodium phosphate, pH7.5, 0.1M NaCl, 1mM EDTA was mixed and

added similarly to the second drop well as negative control. The plate was then sealed by Crystal Clear sealing tape (Greiner) and incubated in a 16°C stable environmental chamber to allow crystallization. The formation of crystal was observed under a light microscope.

2.2.7.3 Optimization of crystallization conditions

The crystallization conditions were optimized by varying the corresponding pH, concentrations of salt and precipitant of the original condition in order to yield bigger and diffracted crystals. Sitting drop method was also performed on 24 well Cryschem Plate. 1 ml crystallization buffer was added to each reservoir. Different protein to buffer ratio, 1μL:1μL, 1μL:2μL and 2μL to 1μL, were also tested.

For the screen yielding good quality diffracting crystal of TM EngB -GDP complex, pH 4.5, 5.5, 6.5 and 7.5 were screened and PEG4000 concentration of 15%, 20%, 25% and 30% were screened in the presence of 0.2M ammonium sulfate. For pH 4.5 and 5.5, 0.2M sodium acetate was used as buffer; for pH 6.5, 0.2M MES was used as buffer; and for pH 7.5, 0.2M Tris-base was used as buffer.

2.2.7.4 X-ray diffraction

A single crystal with appropriate size was picked up by a cryo-loop. The crystal was immersed in mother liquor with cryo-protectant and mounted to a diffractometer cooled by

nitrogen stream at 110 K. The data was collected by an in-house Rigaku MicroMax 007 X-ray generator.

2.2.8 Thermodynamics studies of proteins

2.2.8.1 Preparation of protein sample

Protein stock was concentrated to 0.5mM by ultrafiltration and dialysed for 1000 fold against 20mM sodium phosphate, pH7.5.

2.2.8.2 Guanidine-induced denaturation experiment

0.05mM protein sample in various concentration of guanidine hydrochloride (GdnHCl) was prepared. For each circular dichroism spectrum measurement, 0.5mL of sample was prepared. Protein sample in various concentration of GdnHCl was prepared by mixing protein stock with 20mM sodium phosphate, pH 7.5 buffer and 20mM sodium phosphate, pH7.5 , 8M GdnHCl. Sample was equilibrated at 25°C before measuring CD spectra for at least 16 hours to ensure reaching equilibrium. The mean (3 times) molar ellipticity at 222nm was measured at 25°C by 1mm path length cuvette using JASCO J810 spectropolarimeter equipped with a Peltier-type temperature control unit. For 0M 3.6M and 7.2M GdnHCl, full spectra (260nm to 190nm) were measured to observe the denaturation. For the GdnHCl concentrations in between, 225nm to 215nm were measured. Molar ellipticity at 222nm from 0M to 7.2M were extracted and the data were fitted by non-linear

regression to a two-state model using $Y_{obs} = \{ (Y_n + M_n [D]) + (Y_u + M_u [D]) e^{-G(D)/RT} \} / (1+ e^{-G(D)/RT})$, where Y_{obs} is the observed mean molar ellipticity at 222nm; Y_n and M_n are y-intercept and slope of the linear baseline before the transition; Y_u and M_u are the y-intercept and slope of the linear baseline after the transition; R is the gaseous constant; T is the temperature in Kelvin; [D] is the concentration of GdnHCl; G(D) is the free energy of unfolding at [D].

2.2.8.3 Thermal-induced denaturation experiment

Thermal denaturation was followed by molar ellipticity at 222nm using a JASCO J810 spectropolarimeter equipped with a peltier type temperature control unit. Protein samples were dialyzed in 10mM sodium phosphate buffer, pH7.5 and were degassed before CD measurements. The samples were heated in a 1-mm path length cuvette from 60°C to 110°C at heating rate of 1°C/min.The ellipticity data were then fitted into equation $Y_{obs}= ((a - b)/(1+\exp^{d})) + b$, where $a= Y_n + M_n T$, $b = Y_u + M_u T$, $c= (1/ T) – (1 / Tm)$, and $d= (-H / R) *c$ .The data were fitted for the following parameters: Y_{obs} is the ellipticity at 222nm in millidegrees; Y_n and M_n are y-intercept and slope of the linear baseline before the transition; Y_u and M_u are the y-intercept and slope of the linear baseline after the transition; R is the gaseous constant; T is the temperature in Kelvin; and H is the enthalpy.

2.2.9 Binding Assay to study affinity for ligands

The affinity for the GDP analogue mant-GDP and GDP of the protein TM EngB was estimated.

The affinity was quantified by dissociation constant (Kd).

K_d was defined as, $K_d = [E]_{free}[L]_{free} / [EL]$, at equilibrium

[E]free - Free enzyme concentration
[L]free - Free ligand concentration
[EL] - Enzyme ligand complex concentration

2.2.9.1 Using GDP analogue mant-GDP to detect formation of enzyme-ligand complex (TM EngB-mant-GDP)

Binding of fluorescent MANT-nucleotides to target proteins was investigated through fluorescence resonance energy transfer (FRET) between the MANT-nucleotides (λem at about 440nm) and the tryptophans of the target protein (λex at 285nm) [Remmers *et al.*,1994].

The principle of how mant-GDP work was described in figure 3.7.

2.2.9.2 Basic information of fluorescence spectroscopy

Fluorescence spectroscopy was performed on a Perkin–Elmer LS 50B luminescence spectrophotometer (Perkin–Elmer Analytical Instruments). Binding of fluorescent mant-GDP to TM EngB was investigated through fluorescence resonance energy transfer between the mant-GDP and the tryptophan residues of TM EngB.

3ml samples were placed in a temperature-controlled (25°C) sample holder and the

fluorescence of mant-nucleotides was monitored. The excitation light slight width and emission light slit width were set to 7nm.

Before the titration experiment, the λem (emission wavelength) and λex (excitation wavelength) used to determine the formation of TM1466-mant-GDP complex was determined by making an emission spectrum and an excitation spectrum respectively.

2.2.9.3 Determination of λem and λex

3mL sample containing 0.5μM of TM EngB and 0.4μM mant-GDP in 20mM Tris pH7.5, 0.1M NaCl, 1mM EDTA, 10mM Mg^{2+} was prepared. A sample blank was prepared without protein. Fluorescence emission at 440nm was monitored with excitation light of wavelength from 280-300nm.

λex used in following experiments was determined by looking for the excitation wavelength with largest change in fluorescence ($F_{sample} - F_{blank}$) at 440nm. λex was found to be 285nm. Excitation wavelength at 285nm was used in following experiment.

The fluorescence of the same sample was measured with λex =285nm. λem used in following experiments was determined by looking for the emission wavelength with largest change in fluorescence ($F_{sample} - F_{blank}$). λem was found to be 440nm.

2.2.9.4 Studying ligand affinity by titration with ligand analogue

The titration experiments were done according to Lehoux, I. *et al.*, 2003.

The mant-GDP was titrated from a 0.5mM stock solution to samples containing 3mL of 0.5μM of TM1466 with binding buffer, 20mM Tris pH7.5, 0.1M NaCl, 1mM EDTA, 10mM Mg^{2+}. The dilution was negligible (less than 1%). After incubating for 1 minute, fluorescence emitted at 440nm was recorded with excitation at 285nm.

2.2.10 Pull down experiment to study interacting partner of *E. coli* EngB

E.coli EngB was expressed as dual-tagged protein with SBP (streptavidin binding peptide) and CBP (calmodulin binding peptide) sequential tag at N-terminal end.

2.2.10.1 Preparing protein extracts from *E. coli*

E. coli of selected strain (DH5α) in glycerol stock was inoculated into 5mL LB of starter culture and grown overnight with shaking at 280rpm. 5mL starter culture was inoculated into 500mL LB. Cell growth was monitored by measuring OD600. Cells were harvested when OD600 reached 0.4. Cells were collected by centrifugation at 8000 rpm for 10 min at 4°C (Beckman JA-16.250 rotor). Cell pellets were frozen at -80°C for 30 minutes. 2 grams of pellet was resuspended in 20mL of streptavidin binding buffer (SBB, 20mM Tris pH7.5, 150mM NaCl, 1mM EDTA, 10mM $MgCl_2$). Cells were lysed by sonication with 4s bursts at 1s interval on ice until the suspension became clear.

2.2.10.2 Preparing streptavidin resin

100μL of slurry streptavidin resin (Amersham) was used per pull-down setup. Streptavidin resin was washed with 1mL of streptavidin binding buffer (SBB, 20mM Tris pH7.5, 150mM NaCl, 1mM EDTA, 10mM $MgCl_2$) in 1.5mL centrifuge tube. The resin was centrifuged at 1000x g for 5minutes and 1mL of supernatant was discarded carefully. The washing step was repeated.

2.2.10.3 Binding of dual-tagged *E.coli* EngB to streptavidin resin

100μL of 0.03mM SBP-CBP tagged *E.coli* EngB (EC EngB) was added to 100μL of washed streptavidin resin slurry. 800μL of SBB was added to make up the volume to 1mL. Tagged EngB was allowed to bind to the resin with gentle rotating of the tube at 4°C for 1 hour. The suspension was centrifuged at 1000x g for 5minutes at 4°C to collect the resin. 1mL of supernatant was removed.

The resin was washed by resuspending in SBB and incubated in 4°C for 5minutes to wash away loosely bound protein. The suspension was centrifuged at 1000x g for 5minutes at 4°C to collect the resin. 1mL of supernatant was removed. The washing step was repeated. 100μL of streptavidin resin with tagged EngB bound was prepared.

2.2.10.4 Purifying protein using the prepared streptavidin resin

100μL of tagged EngB-bound streptavidin resin was used per pull-down assay. The prepared resin was added to 10mL of *E.coli* protein lysate described in 2.2.10.1 in a 15mL centrifuge tube. Protein was allowed to bind to the resin by incubating the suspension in 4°C with gentle rotation for 1 hour. 2 control experiments were set up in parallel, one with SBB instead of *E.coli* protein lysate, another with washed streptavidin resin without tagged EngB bound instead of tagged EngB-bound resin.

After incubation, the resin was centrifuged at 1000x g for 15 minutes and 10mL of supernatant was collected as flowthrough in streptavidin affinity binding step.

The resin was washed by resuspending in 1ml SBB and incubated in 4°C for 5minutes to wash away loosely bound protein. The suspension was centrifuged at 1000x g for 5minutes at 4°C to collect the resin. 1mL of supernatant was collected as washing in streptavidin affinity binding step. The washing step was repeated.

Protein on the streptavidin resin was eluted by resuspending the resin in 0.5mL streptavidin elution buffer (SEB, 20mM Tris pH7.5, 2mM biotin, 150mM NaCl, 1mM EDTA, 10mM $MgCl_2$). The suspension was incubated in 4°C for 5minutes to elute streptavdin bound protein including tagged EngB. The suspension was centrifuged at 1000x g for 15minutes at 4°C to collect the resin. The supernatant was collected as biotin eluted protein. The

elution step was repeated. A total of 1mL biotin eluted protein was collected. 50μL aliquot of the eluted protein was stored up for SDS-PAGE analysis.

2.2.10.5 Preparing calmodulin resin

100μL of slurry calmodulin resin (Amersham) was used per pull-down setup. Calmodulin resin was washed with 1mL of calmodulin binding buffer (CBB, 20mM Tris pH7.5, 150mM NaCl, 2mM $CaCl_2$, 10mM $MgCl_2$) in 1.5mL centrifuge tube. The resin was centrifuged at 1000x g for 5minutes and 1mL of supernatant was discarded carefully. The washing step was repeated.

2.2.10.6 Binding of dual-tagged *E.coli* EngB to calmodulin resin

900μL of biotin eluted protein solution was added to 100μL of washed calmodulin resin slurry. 100μL of 20mM $CaCl_2$ was added to the suspension to supplement the calcium ions required for calmoduin binding with CBP . Tagged EngB was allowed to bind to the resin with gentle rotating of the tube at 4°C for 1 hour. The suspension was centrifuged at 1000x g for 5minutes at 4°C to collect the resin. 1mL of supernatant was removed.

The resin was washed by resuspending in 1ml CBB and incubated in 4°C for 5minutes to wash away loosely bound protein. The suspension was centrifuged at 1000x g for 5minutes at 4°C to collect the resin. 1mL of supernatant was collected as washing in calmodulin

affinity binding step. The washing step was repeated.

Protein on the calmodulin resin was eluted by resuspending the resin in 0.1mL streptavidin elution buffer (SEB, 20mM Tris pH7.5, 150mM NaCl, 10mM EDTA). The suspension was incubated in 4°C for 15minutes to elute streptavdin bound protein including tagged EngB. The suspension was centrifuged at 1000x g for 5minutes at 4°C to collect the resin. The supernatant was collected as EDTA eluted protein. The elution step was repeated for 3 times. A total of 4 batches of eluted protein were collected. The eluted protein was stored up for SDS-PAGE analysis.

2.2.10.7 Analysis of dual-tag affinity purified protein

The flow-through, washing and elution in each step with the control experiments were analyzed by 15% SDS-PAGE. To detect the protein in SDS-PAGE, silver stain was used to stain the acrylamide gel.

2.2.11 Silver staining of acrylamide gel

Silver staining was performed as described in the Amersham PlusOne silver staining kit for protein.

2.2.11.1 Silver staining reagents

Fixing solution

Ethanol 75 ml, Glacial acetic acid 25 ml, Water to 250 ml

Sensitizing solution

Ethanol 75 ml, Sodium thiosulphate (5% w/v) 10 ml, Sodium acetate (17 g), add water to 250 ml, before use: Add 1.25 ml glutardialdehyde (25% w/v)

Silver solution

Silver nitrate solution (2.5% w/v) 25 ml, add water to 250 ml, before use: Add 0.1 ml formaldehyde (37% w/v)

Developing solution

Sodium carbonate (6.25 g), add water to 250 ml, stir vigorously to dissolve the sodium carbonate; before use: Add 0.2 ml formaldehyde (37% w/v)

Stop solution

EDTA-Na_2•$2H_2O$ (3.65 g), add water to 250 ml

Washing solution

Water

2.2.11.2 Silver staining procedures

100mL of each solution was used for 7cm width SDS-PAGE. Soak the gel in fixing solution for 60 minutes (or up to 24 hours). Remove the solution. Add sensitizing solution and leave shaking for 60 minutes (or up to 24 hours). Remove the sensitizing solution. Add distilled water and wash four times for 15 minutes each time. Add silver solution and leave shaking for 60 minutes. Remove the silver solution. Wash four times in distilled water for

one minute each time. Add developing solution and leave shaking for 4-6 minutes. Transfer the gel to stopping solution when the bands / spots have reached desired intensity. Leave shaking for 60 minutes.

Chapter Three

3 Structure determination of *T. maritima* EngB by X-ray crystallography

3.1 Introduction

To deepen our understanding towards structural function of EngB, we aimed to determine the crystal structure of *T. maritima* EngB (TM EngB) by X-ray crystallography. Chapter 3 will focus on the structure determination of the apo- and GDP-bound forms of TM EngB, while chapter 4 will focus on detailed analyses of structure models determined. TM EngB was expressed and purified in *E.coli*. To study the structure-function of EngB, obtaining purified EngB is the first step. Purified EngB is subjected to crystallization trials and high quality crystals produced will proceed to X-ray diffraction experiment. With previous *B. subtilis* EngB structure model (PDB code: 1SUI), structure model of TM EngB can be built by molecular replacement. High resolution crystal structure of EngB will help us to understand how structure of EngB is related to its function.

3.2 Generation of *T. maritima* EngB expression construct

To obtain TM EngB protein, a construct for expressing TM EngB with N-terminal fusion maltose binding protein was made. DNA fragment encoding the full length of TM EngB was obtained by PCR described in section 2.2.2.1. The PCR product was analyzed

by 1% TAE agarose gel for its expected size of 600bp (section 2.2.2.2), and was gel-purified and recovered by Viogene Gel-M kit (section 2.2.2.3). The recovered DNA and pRSETA-MBP were cleaved by *Bam*HI and *Eco*RI for an hour (section 2.2.2.4), gel-purified and recovered by Gel-M kit. The recovered insert DNA was then ligated to the cleaved pRSETA-MBP (section 2.2.2.5). The ligation products were directly transformed to competent *E. coli* DH5α (section 2.2.2.6). The transformed bacteria were incubated at 37°C for 14 to 16 hours. Plasmid DNA was isolated from the bacterial colonies (section 2.2.2.8) and those containing DNA insert were screened (section 2.2.2.9) and sequenced (section 2.2.2.10). Figure 3.1 shows the schematic diagram of the expression construct.

3.3 Expression and purification of TM EngB

To express TM EngB in *E. coli*, plasmid pRSETA-MBP-TM1466 was transformed into *E. coli* BL21(DE3)pLysS. Fusion MBP-TM1466 fusion protein was produced by expressing the protein in *E. coli* as described in (section 2.2.3). The cells were harvested and lysed. Proteins in the cells were analyzed by SDS-PAGE (figure 3.2). MBP-TM1466 (65kDA) was found to be over-expressed with a major band shown in SDS-PAGE analysis.

MBP-TM EngB fusion protein was purified by SP-ion exchange chromatography as described in section 2.2.4.1. The MBP tag was cleaved off as described in section 2.2.4.2. After dialysis in 20mM Tris, 0.1M NaCl, TM EngB was further purified by Heparin

affinity chromatography as described in section 2.2.4.3, and then by gel filtration chromatography as described in section 2.2.4.4. The elution profile (figure 3.3) showed a single peak which suggests a single population of protein was found. Protein profile after each purification step can be found in figure 3.4).

3.4 TM EngB was crystallized with freshly purified TM EngB

Preliminary sparse matrix screenings with Hampton Index and Crystal screen 1 and 2 were set up as described as section 2.2.7. Crystals were obtained in various conditions containing PEG4000 as precipitants at pH 4.5 to 7.5 with ammonium salt. Crystallization was optimized by varying pH, concentration of salt and precipitant in corresponding conditions.

Crystals obtained were screened by carrying out diffraction experiment with in-house x-ray source as described in section 2.2.7.4. The best quality crystal with highest resolution obtained was crystallized in 0.2M sodium acetate pH5.5, 30% PEG4000, 0.2M ammonium sulfate. Cryo-protection was achieved by soaking the crystals in mother liquor with 15% (v/v) PEG400 for 10 seconds. The crystal was then loop-mounted and transferred into the cryo-stream of nitrogen at 110 K. X-ray diffraction data were collected on an R-AXIS IV++ imaging-plate system using a rotating copper-anode X-ray source (Rigaku MicroMax-007 with VariMax optics). A total of 180 images were collected at 0.5°

oscillations, with an exposure time of 10 min and a crystal-to-detector distance of 100 mm.

3.5 Data processing of diffraction data and structure refinement of TM EngB

The diffraction data were indexed, integrated, scaled and merged using the programs MOSFLM, SCALA and TRUNCATE from the CCP4 suite (Collaborative Computational Project, Number 4, 1994; Table 3.5). The crystals belong to the orthorhombic space group P212121, with unit-cell parameters a=62.4777Å, b=69.8361Å, c=94.7337Å. With two protein molecules per asymmetric unit, the Matthews coefficient of the crystal is 2.2$Å^3Da^{-1}$ and the solvent content is 44.8% [Matthews, B.W., 1968].

Molecular replacement was performed with the program MOLREP (Collaborative Computational Project, Number 4, 1994) using all data up to resolution range 3Å. A modified model of the apo form YsxC protein (PDB code 1sul; Ruzheinikov, S.N., *et al.*, 2003), which shares 40% sequence identity with TM EngB, was used as the search model. The model was modified by Align4mr [Cheung, Y.Y. *et al.*, 2004] which is a PERL script using sequence alignment to prepare search model. R factor and the correlation factor were 0.614 and 0.209 respectively for the best search result.

5% of the reflection data was picked as test set and did not involve in later refinement. Initial model was refined using CNS [Brünger, A. T. *et al.*, 1998] rigid body refinement and simulated annealing. R factor and R free dropped to 0.4445 and 0.4789 respectively.

Further refinement was done by Xfit [McRee. D.E., 1999] for real space refinement and CNS for energy minimization and b-factor refinement. From the electron density map generated by Xfit, there was a GDP molecule bound on the each of the protein molecule in the model. A GDP molecule was added to the structure model on each protein monomer. Water molecules were added to the model until R_{free} had been lowered to 0.30 by CNS. Further refinement was done using Coot [Emsley, P. *et al.*, 2004] for real-space refinement and CCP4 refmac which include riding hydrogen into refinement. Model was further assessed by WHAT IF [Vriend, G., 1990; Hooft R.W.W, *et al.*, 1996]. Statistics of final structure model was listed as table 3.6.

3.6 Apo-form TM EngB was obtained by unfolding and refolding

It is surprising that GDP molecules are observed in the crystal structure of the purified TM EngB protein. Exogenous GDP was not added in steps of chromatographic purification or dialysis. Therefore, GDP bound on TM EngB was originated endogenously from *E. coli*, which bound to TM EngB once the protein was expressed and was retained after steps of chromatographic purification or dialysis, it was not removed from the protein. The GDP was tightly bound on the protein. High affinity of the protein towards its ligand GDP is demonstrated and it will be further discussed in Chapter 5.

As the GDP molecule is tightly bound to the TM EngB protein, simple dialysis or

desalting will not serve the work to remove it. To remove the bound GDP on TM EngB, dialysis of TM EngB in denatured state was carried out. As shown in chapter 6, TM EngB was capable to refold after chemical unfolding by 7.2M guanidine hydrochloride (GdnHCl).

To unfold TM EngB, purified TM EngB was first concentrated to 1mg/mL. It was then unfolded by dialysis in 100mL 7M GdnHCl, 20mM Tris, pH 7.5, per 1mL of protein solution at 4°C for 4 hours. The dialysis was repeated once with fresh denaturing buffer after the first dialysis to ensure complete exchange of denaturing buffer and removal of bound GDP. TM EngB was refolded by dialyzing in 1L 0.1M NaCl, 20mM Tris, pH 7.5, per 1mL of protein solution at 4°C for 4 hours twice. Refolded protein was subjected to CD spectrum measurement (section 2.2.8) to confirm the refolding of protein.

3.7 Crystallization of apo-form TM EngB

To obtain diffracting crystals of apo-form TM EngB, preliminary sparse matrix screenings with Hampton Index and Crystal screen 1 and 2 were set up as described as section 2.2.7. Crystals obtained in sparse matrix screen were screened by carrying out diffraction experiment with in-house x-ray source as described in section 2.2.7.4. The best quality crystal with highest resolution obtained was crystallized in 0.1 M bis-tris pH 5.5, 2.0 M ammonium sulfate.

Cryo-protection was achieved by soaking the crystals in mother liquor with 30% (v/v) glycerol for 10 seconds. The crystal was then loop-mounted and transferred into the cryo-stream of nitrogen at 110 K. X-ray diffraction data were collected on an R-AXIS IV++ imaging-plate system using a rotating copper-anode X-ray source (Rigaku MicroMax-007 with VariMax optics). A total of 360 images were collected at 0.5° oscillations, with an exposure time of 10 min and a crystal-to-detector distance of 150 mm.

3.8 Data processing of diffraction data and structure refinement of apo-form TM EngB

The diffraction data were indexed, integrated, scaled and merged using the programs MOSFLM, SCALA and TRUNCATE from the CCP4 suite (Collaborative Computational Project, Number 4, 1994; Table 3.5). The crystals belong to the orthorhombic space group C2221, with unit-cell parameters a=69.7173Å, b=97.9882Å, c=54.8474Å. With two protein molecules per asymmetric unit, the Matthews coefficient of the crystal is 2.04 $Å^3Da^{-1}$ and the solvent content is 39.64 % [Matthews, B.W., 1968].

5% of the reflection data was picked as test set and did not involve in later refinement. Initial model was built using TM EngB-GDP bound form. It was refined using CNS [Brünger, A. T. *et al.*, 1998] rigid body refinement and simulated annealing. R factor and R free dropped to 0.2625 and 0.2918 respectively.

A sulfate molecule was added to P-loop on the protein monomer. Water molecules were added to the model until R_{free} had been lowered to 0.30 by CNS. Further refinement was done by COOT [Emsley, P. *et al.*, 2004] for real space refinement, CNS and CCP4 refmac for energy minimization and b-factor refinement. Statistics of final structure model was listed as table 3.6.

3.9 Producing EngB-GDP complex crystal from apo-from EngB

To demonstrate that apo-form EngB was functional and capable to bind GDP molecule, crystal screening of EngB / GDP complex crystal was repeated using apo-form EngB with exogenous GDP purchased from Sigma. Apo-from EngB was prepared to10mg/mL in 10mM Tris, pH 7.5, 0.1M NaCl. 1mM GDP solution prepared in 10mM Tris, pH7.5 was added to apo-form EngB solution in 1:1 volume ratio. Griding crystal screen was made with pH varing from 4.5-7.5 at 1 pH interval, PEG4000 ranging from 15% to 30% at 5% interval and with 0.2M ammonium sulfate as another precipitant as described in chapter 2.2.7.3. Good quality diffracting crystals were obtained at 0.2M sodium acetate, pH 5.5, 30% PEG4000 and 0.2M ammonium sulfate. Diffracting quality crystals were also obtained in this condition using EngB protein without refolding. Crystals obtained was proceeded to diffraction experiment. Cryo-protection was achieved by soaking the crystals in mother liquor with 10% (v/v) glycerol for 10 seconds. The crystal was then

loop-mounted and transferred into the cryo-stream of nitrogen at 110 K. X-ray diffraction data were collected on an R-AXIS IV++ imaging-plate system using a rotating copper-anode X-ray source (Rigaku MicroMax-007 with VariMax optics). A total of 180 images were collected at 0.5° oscillations, with an exposure time of 10 min and a crystal-to-detector distance of 100 mm. Diffraction data has been summarized in Table 3.5a.

The crystal belongs to orthorhombic space group P212121, the same as the previous TM EngB-GDP complex, with unit-cell parameters a=62.6702 Å, b=69.7137 Å, c=94.8355Å, slightly deviated from the previous one.

Molecular replacement was performed with the program MOLREP (Collaborative Computational Project, Number 4, 1994) using all data up to resolution range 3Å. Structural model of previous TM EngB / GDP complex was used as the search model. Initial model was refined using CNS [Brünger, A. T. *et al.*, 1998] rigid body refinement and simulated annealing R factor after annealing was 0.2701 and R_{free} was 0.2935. Single round of energy minimization and b-factor refinement by CNS further brought the R factor to 0.2646 and R_{free} to 0.2916. Statistics for the preliminary refinement statistics was summarized in Table 3.5b.

It was successful to prepare co-crystal of TM EngB / GDP complex from apo-form EngB protein. By inspecting the structural model with electron density map generated by

REFMAC, two monomers of EngB were observed in each asymmetric unit and on each EngB, a GDP molecule was bound. No major structural differences were observed between the previous TM EngB / GDP complex structure and this crystal structure prepared from apo-form EngB.

3.10 TM EngB is a monomer in solution

2 monomers were observed in the TM EngB/GDP complex structure. To determine if TM EngB is a dimmer in native state, native molecular weight of TM EngB was measured by light scattering in chromatography mode as described in section 2.2.12. Binding of guanine nucleotide may change the stoichiometry of TM EngB, therefore the native molecular weight of TM EngB in apoenzyme, in GDP-bound and GTP- bound state have to be determined. Samples of TM EngB were prepared as apoenzyme, GDP-bound enzyme, and GMPPNP-bound enzyme. GMPPNP was used as an unhydrolysable analogue of GTP. Samples were pre-fractionated by size exclusion chromatography to remove ambiguities that might result from a batch mode experiment, where the measured quantities are averaged over the distribution of masses and sizes present in the sample. Moreover, a chromatography mode experiment makes it possible to completely quantify the molar mass and size distribution for the sample. Native molecular weight of each sample was determined. Apo TM EngB showed a molecular weight of 28.0kDa, TM EngB with GDP

showed a molecular weight of 23.4kDa, TM EngB with GMPPNP showed a molecular weight of 24.6kDa (figure 3.7). Theoretical molecular weight of TM EngB should be 22.5kDa. The deviation of the measured molecular weight may be due to non-specific temporal aggregation of TM EngB, increasing the average molar mass of protein pass through, as each sample does pass through the size exclusion column as a single major population. From the results, TM EngB does exist as a monomer in solution and this is not affected by the binding of guanine nucleotide.

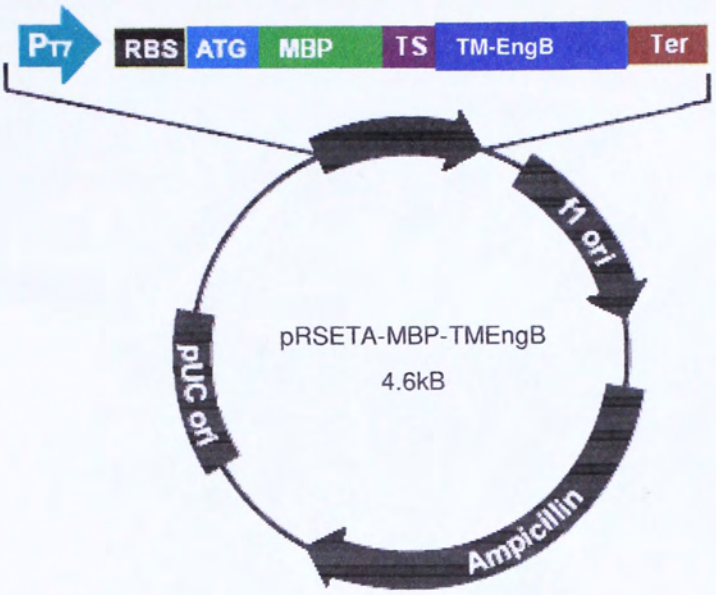
3.11 Summary of Chapter Three

Crystal structures of TM EngB of GDP-bound and apo form have been solved. TM EngB protein retained a bound GDP molecule after steps of purification, and it has been demonstrated that the GDP molecule can be removed by unfolding the protein followed by refolding to obtain apo-form EngB. Secondary structure composition of the refolded apo-form EngB does not change greatly after refolding suggested by circular dichroism spectrometry. The refolded EngB does not lose its function of binding guanine nucleotide as suggested by crystallography study, where GDP bound TM EngB co-crystal can be obtained from the refolded apo-enzyme by adding exogenous GDP to it. Therefore the refolded apo-form TM EngB is in native state.

In the structure of TM EngB /GDP complex, two monomers of EngB are found in

each asymmetric unit, while that of the apoenzyme contains only one monomer in each unit. To demonstrate the effect of binding of guanine nucleotide ligand to the stoichiometric state of TM EngB, light scattering technique is employed to determine the native molecular weights and sizes of the protein. All three samples, apo form TM EngB, apoenzyme mixed with GDP and apoenzyme mixed with GMPPNP, showed a native molecular weight of 22-25kDa. It matches the monomer molecular weight 22kDa of TM EngB. Therefore, TM EngB exists as monomer in solution and binding of guanine nucleotide does not change its stoichiometry.

Figure 3.1 Map of expression construct pRSETA-MBP-TM EngB



P_{T7} : T7 Promoter

RBS: Ribosome binding site

ATG: Start codon

MBP: Maltose binding protein

TS: Thrombin clevage site, encoded for peptide sequence LVPRGS

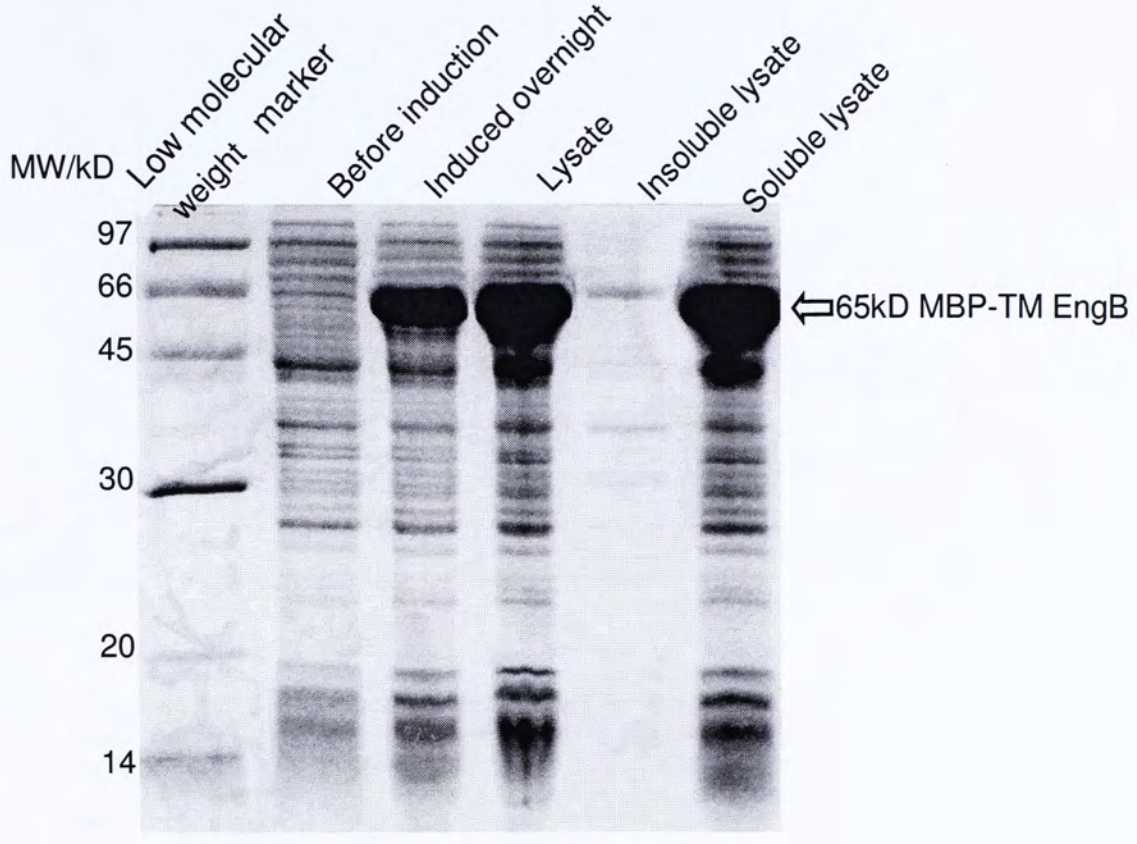
Ter: T7 Terminator

pUC ori: pUC origin of replication

f1 ori: f1 origin of replication

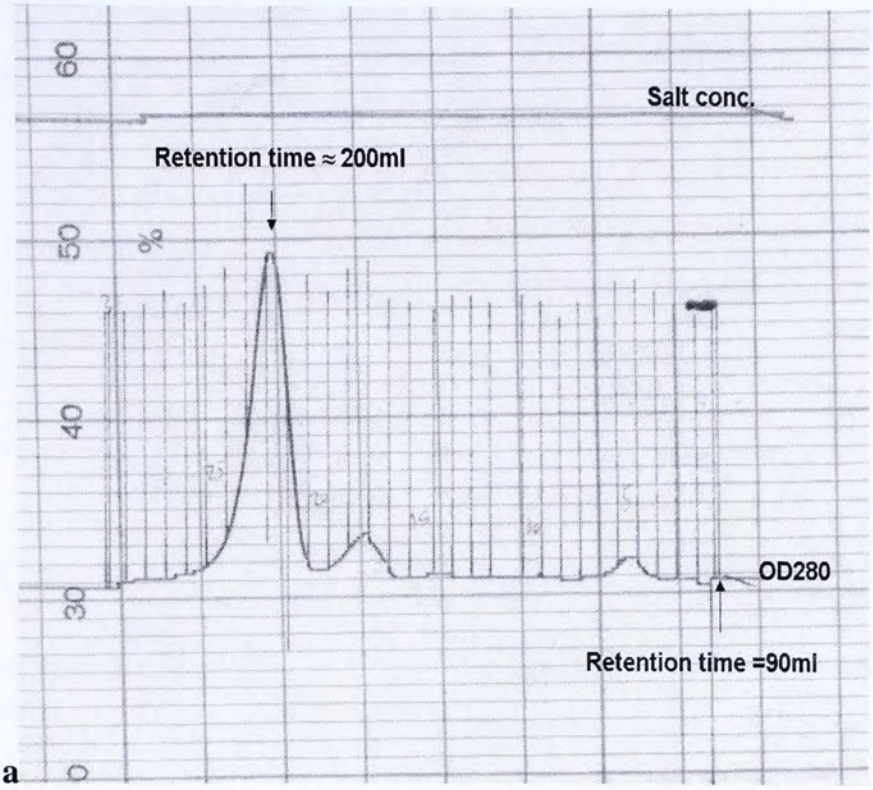
Ampicillin: ampicillin resistant gene

Figure 3.2 *T. maritima* EngB was expressed as a fusion protein with N-terminal maltose binding protein

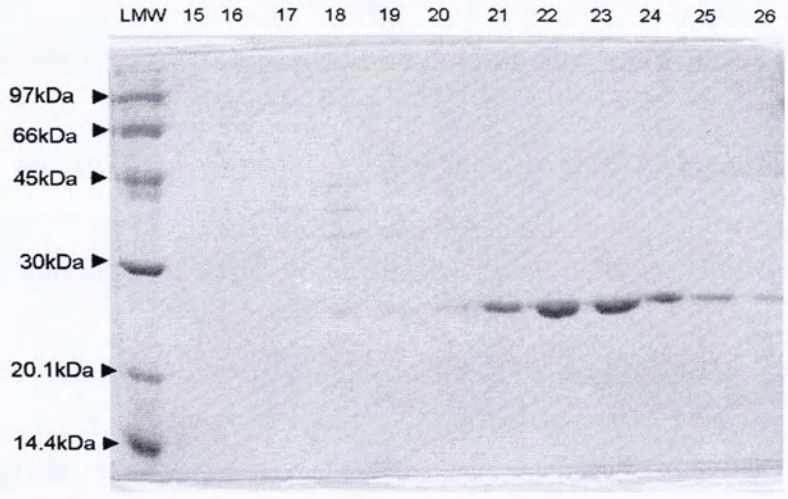


T. Maritima EngB with N-terminal fusion maltose binding protein expressed in soluble fraction using *E.coli* BL21 (DE3) pLysS. Protein in each phase of expression and cell lysis was analyzed by SDS-PAGE and stained by Coomassie Blue. After the expression culture reached OD600=0.4, IPTG was added to 0.4mM concentration to induce expression of the fusion protein. 16 hours after induction, cells were collected by centrifugation and were disrupted by sonication. Lysate was centrifuged and the supernatant which contains the soluble fraction of protein was collected for further purification.

Figure 3.3 *T. maritima* EngB was purified



3.3 a

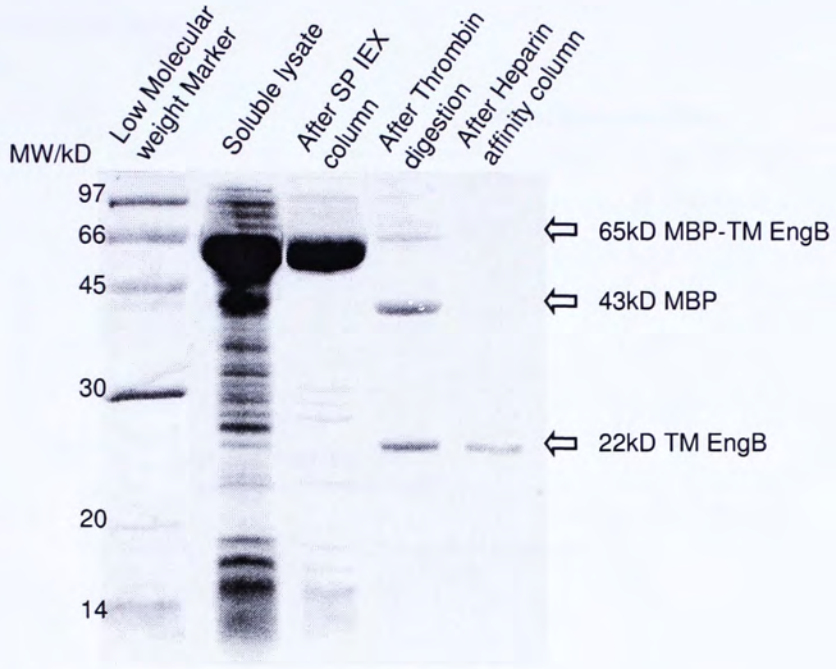


3.3b

(a) The elution profile of TM EngB with gel filtration column. OD280 was measuring the amount of protein eluted. Single population of TM EngB was obtained from fraction 20-26 as a single peak in the elution profile.

(b) Fractions 15-26 in gel filtration chromatography were analyzed by SDS-PAGE. Fractions 20-26 were collected and pooled together as purified TM EngB.

Figure 3.4 *T. maritima* EngB purification profile



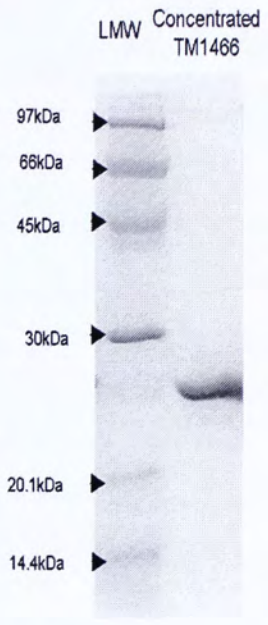
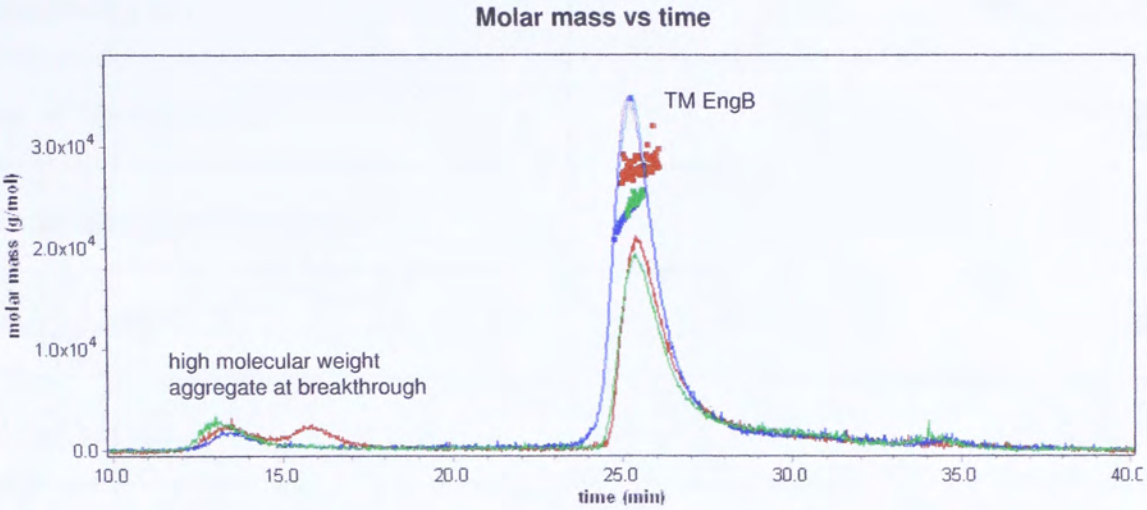


Figure 3.4 A **Figure 3.4 B**

A. *T. Maritima* EngB was purified by column chromatography. Protein profile after each step was analyzed by SDS-PAGE and stained with Coomassie Blue. Fusion MBP-TM EngB (65kD) was purified by SP ion exchange column. Fusion protein was cleaved by protease thrombin, which gave non-fusion TM EngB(22kD). TM EngB was purified by heparin affinity column.

B. TM EngB purified after gel filtration column and concentrated to 0.3mg/mL. Purified TM EngB was analyzed by SDS-PAGE stained by Coomassie Blue. The protein was pure as seen from Coomassie Blue stained SDS-PAGE.

Figure 3.5 Light scattering study of TM EngB proved it to be a monomer in native solution form



Lines represent the reflective index measured; dots represent the calculated molar mass of fractions pass through the detector. Molar mass of apo TM EngB (red), TM EngB with GDP (blue) and TM EngB with GMPPNP (green). The reflective index represents the relative amount of protein pass through the detector. From the results, the retention times for all three samples are similar. The average molar mass of apo TM EngB was found to be 28.0kDa, that of TM EngB bound with GDP was 23.4kDa and that of TM EngB bound with GMPPNP was 24.6kDa. Theoretical molecular weight of TM EngB should be 22.5kDa. TM EngB from all 3 samples exists as a monomer in solution.

Table 3.1 Data-processing statistics for *TM* EngB GDP complex

Values in parentheses are for the highest resolution shell (1.84-1.75).

Resolution (Å)	**39.19 - 1.75**
No. of observations	**275194**
No. of unique reflections	**40506**
Multiplicity	**6.8(6.2)**
Completeness	**95.2(99.1)**
$\mathbf{R_{merge}}^{+}$	**0.065(0.367)**
Mean I/ σ(I)	**21.3(3.9)**

$^{+} R_{merge} = \sum_{hkl} |I - (I)| / \sum_{hkl} I$

Table 3.2 Crystallographic Data and Refinement Statistics for TM EngB GDP complex

Crystal Information	
space group	P212121
unit cell dimensions (Å)	a=62.4777, b=69.8361, c=94.7337
Refinement Statistics	
resolution range (Å)	56.25-1.75
R_{cryst}	0.2162
R_{free}	0.2531
model	
no. of protein atoms	2921
no. of water atoms	197
no. of GDP molecule	2
average B factors (Å^2)	20.047
rms deviation from ideality	
bonds (Å)	0.30
bond angles (deg)	2.28
coordinate error (Å)	
estimated from Rfree	0.133
estimated from maximum likelihood	0.082
geometry, Ramachandran plot analysis	
most favored region (%)	90.1
additional allowed region (%)	9.8
generously allowed and disallowed regions (%)	0
missing residues	A:1-2, 82-83,93-94; B:1-4, 44-52,81-83, 193-194

$R_{cryst} = \sum | |Fo| - |Fc| | / \sum |Fo|$, where Fo and Fc are the observed and calculated amplitudes, respectively. The free set contains 2015 (5% of total) reflections.

Table 3.3 Data-processing statistics for apo form TM EngB

Values in parentheses are for the highest resolution shell (2.42-2.3)

Resolution (Å)	**48.97 - 2.3**
No. of observations	**56299 (8530)**
No. of unique reflections	**8326 (1236)**
Multiplicity	**6.8 (6.9)**
Completeness	**96 (100)**
R_{merge} [+]	**0.102 (0.269)**
Mean I/ σ(I)	**19.4 (4.80)**

[+] $R_{merge} = \sum_{hkl} |I - (I)| / \sum_{hkl} I$

Table 3.4 Crystallographic Data and Refinement Statistics for apo-form TM EngB

Crystal Information	
space group	C2221
unit cell dimensions (Å)	a=69.717 b=97.988 c=54.847
Refinement Statistics	
resolution range (Å)	56.80 - 2.30
R_{cryst}	0.2531
R_{free}	0.2760
model	
no. of protein atoms	1450
no. of water molecule	38
no. of sulfate molecule	1
average B factors (Å^2)	26.896
rms deviation from ideality	
bonds (Å)	0.021
bond angles (deg)	1.690
coordinate error (Å)	
estimated from Rfree	0.293
estimated from maximum likelihood	0.179
geometry, Ramachandran plot analysis	
most favored region (%)	88.1
additional allowed region (%)	11.9
generously allowed region (%)	0
missing residues	46-56,78-83,193-194

$R_{cryst} = \sum | |Fo| - |Fc| | / \sum |Fo|$, where Fo and Fc are the observed and calculated amplitudes, respectively. The free set contains 392 (5% of total) reflections.

Table 3.5a Data-processing statistics for TM EngB GDP complex (apo-form EngB co-crystal with GDP)

Values in parentheses are for the highest resolution shell (2.00-1.90).

Resolution (Å)	**23.71-1.90**
No. of observations	**118351**
No. of unique reflections	**33459**
Multiplicity	**3.5(3.5)**
Completeness	**99.9(100)**
R_{merge} [+]	**0.070(0.366)**
Mean I/ σ(I)	**15.9(3.2)**

[+] $R_{merge} = \sum_{hkl} |I - (I)| / \sum_{hkl} I$

Table 3.5b Crystallographic Data and Preliminary Refinement Statistics for TM EngB-GDP complex (apo-form EngB co-crystal with GDP)

Crystal Information	
space group	P212121
unit cell dimensions (Å)	a=62.6702 b=69.7137 c=94.8355
Refinement Statistics	
resolution range (Å)	23.71 – 1.90
R_{cryst}	0.2646
R_{free}	0.2916
model	
no. of protein atoms	2982
no. of GDP molecules	2

$R_{cryst} = \sum | |Fo| - |Fc| | / \sum |Fo|$, where Fo and Fc are the observed and calculated amplitudes, respectively. The free set contains 1692 (5% of total) reflections

Chapter Four

4 Structural analyses of TM EngB

4.1 Introduction

GTPases act as a major signaling mechanism operating as molecular switches via the hydrolysis of GTP. They are involved in many important functions, e.g., mRNA translation, cell cycling and signal transduction. The small GTPases share a common structural fold, with the central β-sheet flanked by α-helices. Residues involved in nucleotide binding are located in five motifs, G1-G5. The sequence pattern of G1-G5 is used to group the GTPase into subfamilies [Lipe,D.D., *et al.*, 2002]. Upon binding of different nucleotides, structural changes happen in the switch I and switch II regions, in which G2 and G3 is covered respectively. Structural change in those regions may contribute to the functions of the protein. Here we report the crystal structures of *T. maritima* EngB in apo-enzyme form and GDP bound form. They are compared with each other and with EngB from *B. subtilis*. The conformation of the switch I loop in GDP-bound EngB was firstly reported in this work. This may give information for us to track important conformational change which led to the change of on/off state on this protein.

4.2 Overall fold of TM EngB

Two monomers of TM EngB were observed in the crystal structure of TM EngB/GDP

complex. A single monomer of TM EngB was observed in the crystal structure of apo TM EngB.

Secondary structure of both chains in crystal structure of TM EngB/GDP complex and that of apo TM EngB were analyzed (figure 4.1) and summarized in a topology diagram (figure 4.2).

The TM EngB subunit consists of a single polypeptide chain which folds into a single domain. The folding pattern is based on an α/β-type fold, with central β-strands flanked by the α-helices. TM EngB resembles fold of pattern of other TRAFAC class GTPases [Leipe, D.D. *et al.*, 2002]. p21ras, EF-Tu and other Ras-like GTPases share similar folding pattern with TM EngB. Comparing with p21ras, TM EngB contains one additional beta strand at the N-terminus, a total of seven beta strands at the center.

4.3 Mode of nucleotide binding of TM EngB

By analyzing the structure and mode of nucleotide binding of TM EngB, structurally aligned with *Bacillus subtilis* YsxC [Ruzheinikov, S.N., *et al.*, 2004] of the same family, conserved binding motifs of GTP-binding proteins can be recognized (figure 4.1 and 4.3). This helps us to identify important regions to study and analyze the structure of TM EngB.

The G1 region (residues 29-36) covers part of the P-loop (residues 31-38). The P-loop is responsible for the recognition of the α and β- phosphate group of the GDP nucleotide.

The sequence of G1 region in TM EngB is GRSNVGKS, it contains the sequence motif "GxxxxGKS/T" [Leipe, D.D. *et al.*, 2002] of the superfamily GTPase. This region is responsible for forming multiple H-bonds with α and β phosphate groups of the GDP (figure 4.3). In apo form TM EngB crystal structure, the P-loop is bound with a sulfate molecule.

The G2 region (residues 48-60) covers part of the switch I region (residues 45 - 60). In this region, 6/13 residues is conserved between TM EngB and YsxC. In chain A of the TM EngB/ GDP structure, backbone nitrogen on Ser51 (corresponding to Ser53 in YsxC) recognizes the alpha phosphate of the bound nucleotide. Thr59 in YsxC/GMPPNP/Mg^{++} complex is responsible for the specific recognition of magnesium cation in this region [Ruzheinikov, S.N., *et al.*, 2004]. The corresponding residue in TM EngB, Thr57 is conserved. However, without a GTP (or analogue)/Mg^{++} complex structure of TM EngB, function of Thr57 as magnesium recognition site is still unclear. Sidechain of Thr57 in the crystal structure of TM EngB/GDP is pointing towards the solvent environment.

The G3 region (residues 73-76) is part of the switch II (residues 73-101). The sequence DLPG contains the conserved Walker B motif "hhhhDxxG" (h stands for any hydrophobic residues) of GTPase. In YsxC/GMPPNP/Mg^{++}, side chain of Asp75 in the motif is responsible for the specific interaction with the bound magnesium. In TM EngB/GDP complex, the corresponding Asp73 sidechain forms hydrogen bond with the

sidechain nitrogen on Lys35 (figure 4.3). Backbone oxygen of Leu74 in the G3 motif also forms hydrogen bond with Lys35 sidechain nitrogen. Lys 35 is part of the G1 motif, the corresponding residue Lys36 in YsxC/GMPPNP/Mg^{++} is responsible for the recognition of the γ phosphate group of the GMPPNP. Without a γ phosphate group, the hydrogen donor and acceptor on the Lys35 and Asp73 are stabilized by forming hydrogen bond.

The G4 motif in TM EngB is the residues 136-143 with sequence TIVLTKMD. It contains the conserved "N/TKxD" sequence of the GTPase superfamily. Asp143 recognize the two hydrogen donor nitrogen on the guanine ring by forming direct contact with it (figure 4.3). Lys141 in the region form hydrophobic interaction with the guanine ring.

The G5 motif in TM EngB is the residues 172-174 with sequence TSS. In the superfamily of GTPase, the G5 motif involves in nucleotide binding through their mainchain atoms. In the case of TM EngB/GDP complex, backbone nitrogen on Ser174 forms hydrogen bond with guanine oxygen O6 and the sidechain oxygen of it forms hydrogen bond with guanine nitrogen N7, recognizing the guanine ring through 2 hydrogen donor, acceptor matching.

4.4 Structural differences in switch I region between chain A and chain B in crystal structure of TM EngB/GDP complex

In the crystal structure of TM EngB/ GDP complex, there are two monomer of TM

EngB molecule, which are labeled chain A and B. Chain A and chain B are overlaid and aligned together by their Cα atoms (figure 4.4a) with rms deviation of 0.433Å, suggesting the overall fold of the two forms is similar. From the overlaid structures, major structural difference between chain A and chain B is in the switch I region (residues 45-60) .

In chain A, the switch I region forms a well-defined loop. Several interactions between the residues on the switch I loop and the protein core restrict the flexibility of the loop. First, the N-terminal end of switch I loop forms a turn, which is stabilized by the interaction between the backbone nitrogen and oxygen of Ala48 and side chain oxygen and nitrogen of Asn40 respectively. Second, backbone nitrogen Ser51 forms hydrogen bond with alpha phosphate of the bound GDP. Third, backbone oxygen of Thr53 forms hydrogen bond with side chain nitrogen of Asn32.

In chain B, residues 44-52, which cover the N-terminal part of switch I, are missing. It indicates the flexibility in the region. The C-terminal part (residues 53-60) of the switch I of chain B forms defined structure as sidechain of Asn61 recognize the backbone atoms of Gly55 forming a loose turn from residues 55 to 61.

In the YsxC/GDP complex structure, residues 47-59, which cover most part of switch I (residues 46-62), are missing [Ruzheinikov, S.N., *et al.*, 2004]. This suggests the region is flexible. Compare with the TM EngB/GDP complex structure, the whole switch I of chain A is defined and C-terminal end of switch I of chain B is defined. Switch I in TM

EngB/GDP structure does not have any direct contact with protein molecules in crystal lattice, therefore, the observation of a defined switch I region is not due to crystal packing.

To further study the switch I conformation, we compare the crystal structure of TM EngB/GDP chain A with *B.subtilis* YsxC/GMPPNP/Mg^{++} in the switch I region. The interactions between the bound nucleotides, magnesium cation and switch I result in the conformation of switch I in each case. In the structure of YsxC/GMPPNP/Mg^{++}, switch I region, which includes the G2 motif, interact with the bound GMPPNP through the backbone nitrogen of Ser53, Lys58 and Thr59 (figure 4.6). Ser53 interacts with the α phosphate group; Lys58 and Thr59 interact with the γ phosphate group of the bound GMPPNP. Besides, sidechain oxygen of Thr59 interacts with the bound magnesium cation which is positioned side by side with the β and γ phosphate of the bound nucleotide. With all those interactions, switch I is positioned next to the bound GMPPNP and magnesium cation.

On the other hand, there is only one hydrogen bond linking switch I in chain A of TM EngB to the bound GDP. Backbone nitrogen of Ser51 (corresponds to Ser53 in YsxC) interact with the α phosphate group of the GDP. It and the other interactions of switch I bring the N-terminal part of switch I to resemble the conformation found in YsxC/GMPPNP/Mg^{++} complex structure. The C-terminal part of switch I does not form any interaction with the bound ligand. The major cause of difference in conformation of

C-terminal end of switch I between TM EngB/GDP and YsxC/GMPPNP/Mg^{++} complex structures, is the interaction of Lys58 and Thr59 with γ-phosphate of GMPPNP in YsxC complex. The hydrogen bonds between the backbone nitrogen of Lys58 and Thr59 and γ-phosphate of GMPPNP hold the C-terminal part of switch I in close contact with the bound ligand. As a result, conformation of the C-terminal part of switch I of chain A TM EngB/GDP complex is loosely attached on the protein core.

4.5 Structural difference between TM EngB/GDP complex and apo TM EngB

Overall fold of TM EngB/GDP complex and apo TM EngB is similar (figure 4.7). Switch I of apo TM EngB is not detected in the crystal structure. It indicates the high flexibility of the switch when it is ligand free. Besides, residues 78-83 are not detected in the apo form structure. Binding of GDP ligand restrains the flexibility of the region, which is part of the switch II.

4.6 Summary of chapter four

Crystal structure of TM EngB/GDP complex contains 2 independent monomers of TM EngB. Each monomer is a single polypeptide chain folded into a single domain which resembles an α/β-fold with central β-strands flanked by the α-helices, like other TRAFAC class small GTPases. The two polypeptide chains share the same folding pattern,

while major conformation difference is in the switch I region, which is a flexible loop with major structural change upon binding of nucleotides [Ruzheinikov, S.N., *et al.*, 2004].

N-terminal part of Switch I of chain A resembles the conformation of a GTP bound EngB, which is like the GMPPNP-bound YsxC. Ser51 is interacting with the bound GDP α phosphate group, therefore the N-terminal of switch I loop is lying close the bound GDP and α1 helix. The remaining C-terminal part, without the triphophaste, has no anchoring site. It remains a free loop in the solvent space. Conformation of switch I of chain B does not involve the interactions with its bound ligand. Most part (the N-terminal end) of the switch I of chain B is not detected, indicating its flexibility. The remaining C-terminal part is defined; the loop is hold lying besides β3 strand, involving interaction with the β3 strand. In apo TM EngB, the whole switch I is not structured.

By comparing the 3 monomers, it shows that binding of ligand GDP to TM EngB can limit the flexibility of the switch I. Though two conformations are observed in the GDP bound form, binding of GDP is limiting the possibility of loop conformation as we can detect the full switch I loop in TM EngB/GDP chain A structure. It shows the possibility of the switch I loop in GDP bound state partially achieving the conformation of GTP-bound state. This observation contradicts to the previous study in *B. subtilis* YsxC [Ruzheinikov, S.N., *et al.*, 2004]. In that case, the switch I loop in GDP-bound state is completely disordered and is not detected in the crystal structure. From our study, the switch I loop of

EngB, is at least partially ordered, involving more than one conformation possibility of the loop. Therefore, we purpose that change in conformation in the C-terminal part of switch I (residues 53-59 in *T. maritima* EngB) is the key difference causing the change in behavior of EngB functions.

Figure 4.1 Secondary structure comparison of TM EngB in GDP bound and apo form structures

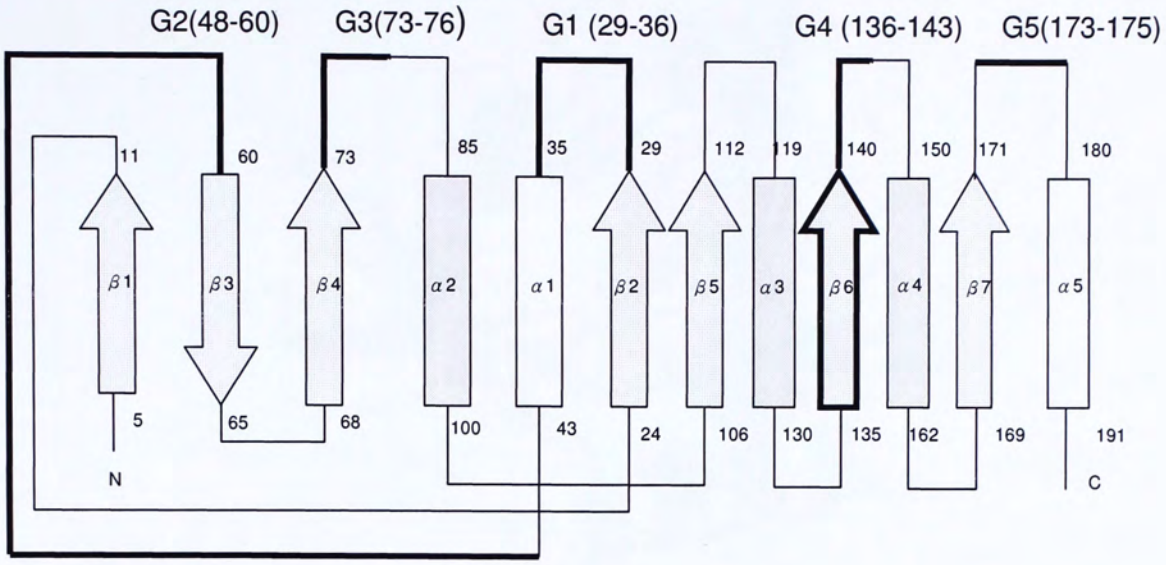
```
                10        20        30        40
        ....|....|....|....|....|....|....|....|
    1   IIRDVELVKVARTPGDYPPPLKGEVAFVGRSNVGKSSLLN 40
GDP-A   --  EEEEEEE             EEEEEE     HHHHHH
GDP-B   ----EEEEEEE             EEEEEE     HHHHHH
APO         EEEEEEE             EEEEEE     HHHHHH
                50        60        70        80
        ....|....|....|....|....|....|....|....|
    41  ALFNRKIAFVSKTPGKTRSINFYLVNSKYYFVDLPGYGYA 80
GDP-A   HHH                 EEEEEE   EEEEE
GDP-B   HHH---------        EEEEEE   EEEEE
APO     HHH  ----------     EEEEEE   EEEEE    ---
                90       100       110       120
        ....|....|....|....|....|....|....|....|
    81  KVSKKERMLWKRLVEDYFKNRWSLQMVFLLVDGRIPPQDS 120
GDP-A    --HHHHHHHHHHHHHHHHH   EEEEEEEEE      HH
GDP-B   ---HHHHHHHHHHHHHHHHH   EEEEEEEEE      HH
APO     ---  HHHHHHHHHHHHHHH     EEEEEEE      HH
                130       140       150       160
        ....|....|....|....|....|....|....|....|
    121 DLMMVEWMKSLNIPFTIVLTKMDKVKMSERAKKLEEHRKV 160
GDP-A   HHHHHHHHHH    EEEEEE GGG  GG HHHHHHHHHHH
GDP-B   HHHHHHHHHH    EEEEEE GGG  GG HHHHHHHHHHH
APO     HHHHHHHHHH    EEEEEE GGG  GG HHHHHHHHHHH
                170       180       190
        ....|....|....|....|....|....|....
    161 FSKYGEYTIIPTSSVTGEGISELLDLISTLLKEN 194
GDP-A   HH      EEE         HHHHHHHHHHHHH--
GDP-B   HH      EEE         HHHHHHHHHHHHH--
Apo     H       EEE         HHHHHHHHHHHHH--

Keys:
GDP-A  TM EngB/GDP complex structure, chain A
GDP-B  TM EngB/GDP complex structure, chain B
APO    TM EngB apo-enzyme structure

H  - 4-helix (alpha-helix)
G  - 3-helix (3/10 helix)
E  - Beta sheet
-  - missing residues
```

Secondary structure of GDP-bound and apo form TM EngB are similar. The conserved regions G1-G5 are underlined and the switch I and switch II regions are highlighted. Conformational change happens in the switch regions upon binding of different guanine nucleotides. Residues in switch I and switch II loop region of the apo form structure are not observed in the apo-enzyme structure, it may indicate higher flexibility in those regions.

Figure 4.2 Topology diagram showing the secondary structure of TM EngB with the conserved G1-G5 domain labeled



The topology of the secondary structure elements of *T. maritima* EngB/GDP. β-strands are represented in arrows and α-helices above and below the plane of the β-sheet are shown as un-shaded and shaded boxes, respectively. The G1-G5 motifs are bolded and labeled with the corresponding residues indicated in brackets.

Figure 4.3a G1, G2, G4 and G5 recognize the bound GDP in TM EngB/GDP complex

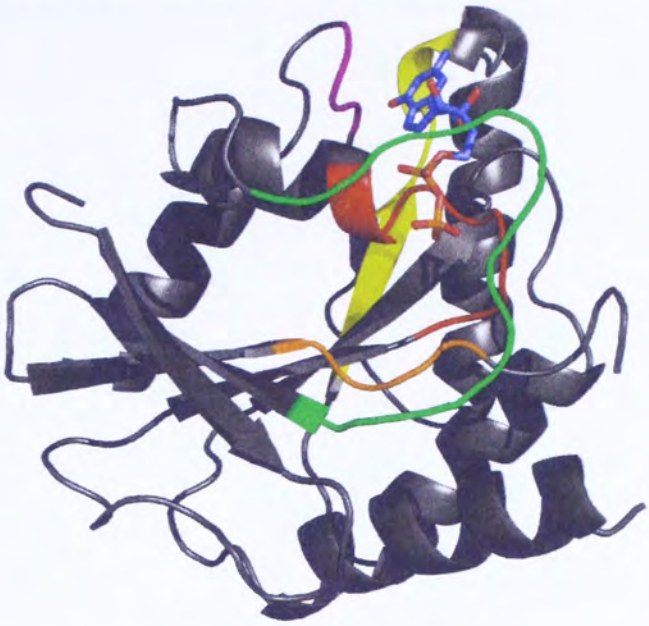
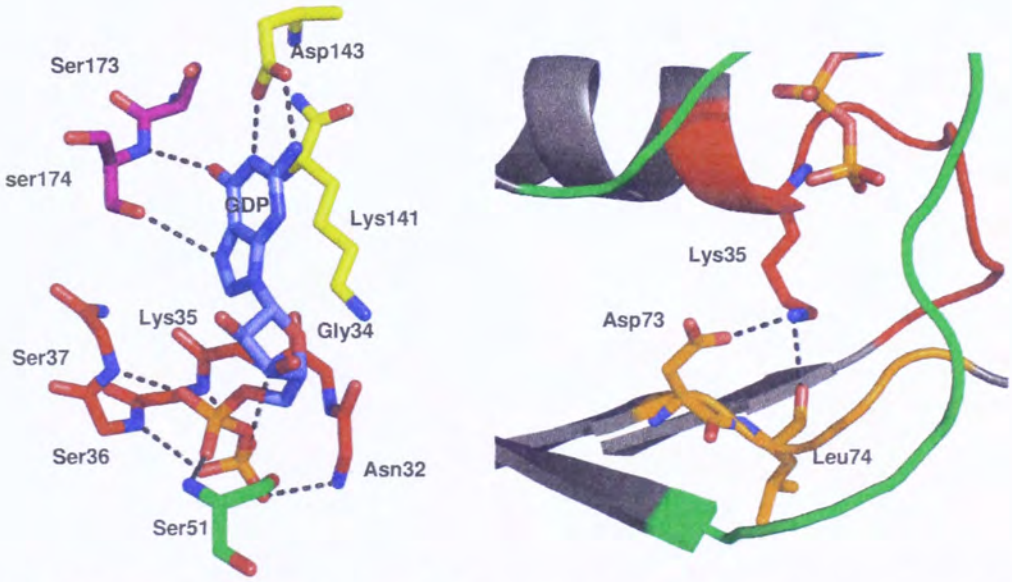


Figure 4.3b (Left) Interactions between TM EngB with the bound GDP
Figure 4.3c (Right) G3 motif in TM EngB/GDP complex does not take part in recognition of the bound GDP



(a) G1 (red), G2(green), G4 (yellow) and G5 (magenta) are involved in the recognition of the bound GDP in TM EngB/GDP complex. (b) The detail interactions with the bound GDP are shown. The interactions with the bound ligand found in chain A and B of TM EngB/GDP complex are the same except Ser51 as most of the G2 motif (green) in chain B was not detected in the structure. Summary of the interactions is listed in table 4.1. (c) G3 (orange) motif is not involved in the recognition of bound GDP. Corresponding residues in G3 motif of *B. subtilisi* /GMPPNP /Mg^{++} complex structure [PDB code: 1SVW] are involved in the recognition of bound magnesium. Without bound magnesium,

Asp73 and Leu74 in the motif are interacting with Lys35.

Figure 4.4a Major differences between chain A and chain B of TM EngB/ GDP complex structures are in the switch I region

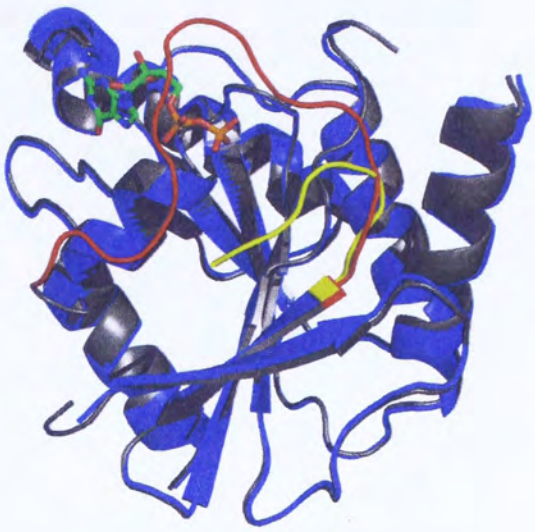
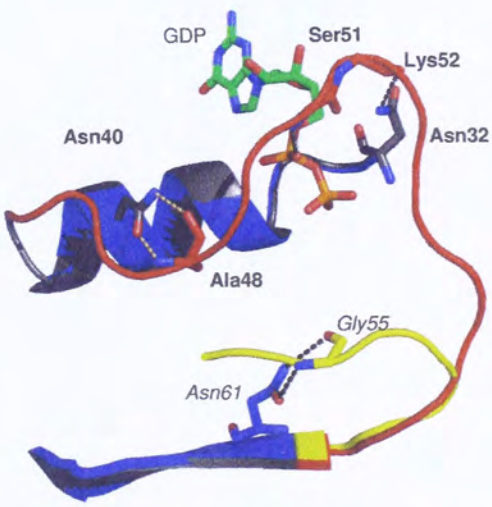


Figure 4.4b Differences between chain A and chain B of TM EngB/ GDP complex structures in switch I region



(a) Switch I of chain A in TM EngB /GDP is colored red, other region in chain A is colored in grey. Switch I of chain B is colored in yellow, other region in chain B is colored in blue. Major differences between the two chains are in the switch I loop region. The N-terminal end of switch I in chain B is not detected. It may indicate the flexibility of switch I in the region. In chain A, the whole switch I is defined and lies adjacent to the bound GDP. (b) Residues involving in interaction in the switch I region is shown in sticks. In chain A, the whole switch I is defined. Backbone nitrogen of Ser51 is interacting with the oxygen on the alpha phosphate of the bound GDP. Ala48 and Lys52 are also involved in defining the conformation of the N-terminal end of switch I loop. These interactions are only found in chain A. In chain B, only the C-terminal end of the switch I loop is observable (from residue 53). The loop is fixed by the interaction between the backbone of Gly55 and

sidechain of Asn61. This interaction is only found in chain B.

Figure 4.5 Structure alignment of TM EngB /GDP complex chain A with *B. subtilis* EngB/GDP complex [PDB code: 1SVI]

Switch I of TM EngB /GDP complex chain A is colored in red; switch II of it is colored in blue; other region of it is colored in grey. Switch I of *B. subtilis* EngB/GDP complex is colored in orange; Switch II of *it* is colored in magenta; other region of it is colored in white.

The overall structures of TM EngB /GDP complex chain A and *B. subtilis* EngB/GDP complex are similar, major differences are found in the switch I (red and orange respectively) and switch II (blue and magenta respectively) region. For the difference in switch I, that in *B. subtilis* EngB/GDP complex is not defined, while that in TM EngB /GDP complex chain A is defined. For the difference in switch II, that in TM EngB /GDP complex (both chains, see figure 4.4a) is not well defined. Though switch II loop region in *B. subtilis* EngB/GDP complex is defined, no specific hydrogen bonds or charge-charge interactions are found in the switch II loop.

Figure 4.6a Switch I of TM EngB/GDP complex resembles the conformation found in *B. subtilis* EngB /GMPPNP /Mg^{++} complex [PDB code: 1SVW]

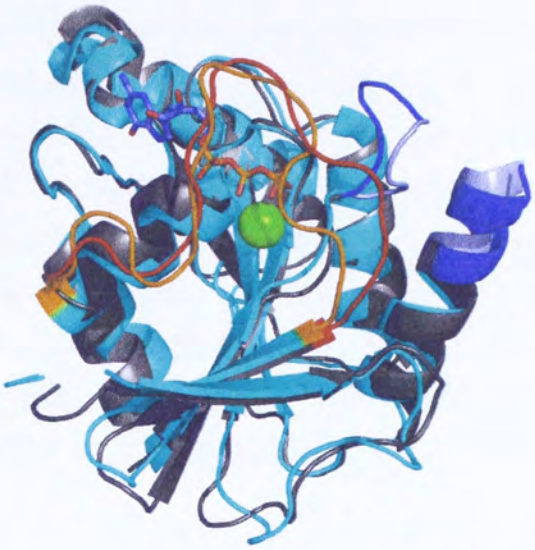
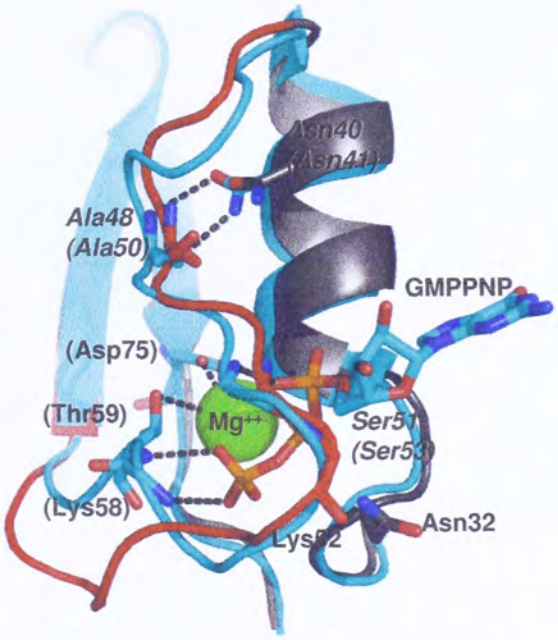


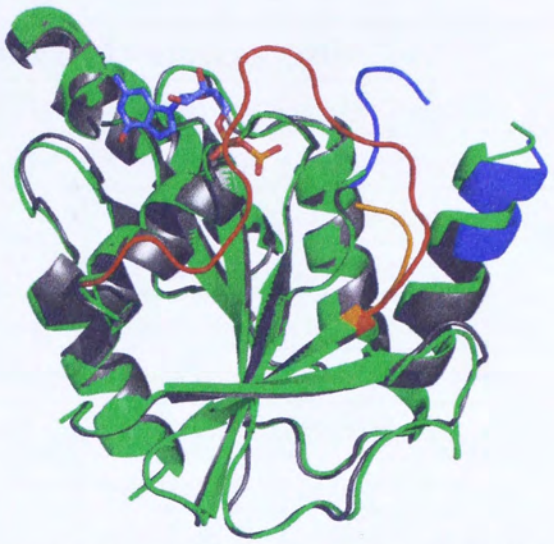
Figure 4.6b Detail conformational differences between TM EngB/GDP complex and *B. subtilis* EngB /GMPPNP /Mg^{++} complex involving the bound ligands



(a) Chain A of TM EngB/GDP complex is aligned with the *B. subtilis* EngB /GMPPNP /Mg^{++} complex [PDB code: 1SVW]. Switch I of TM EngB is colored in red, switch II of it is colored in blue, other region of it is colored in grey. Switch I of *B. subtilis* EngB is colored in orange, switch II of it is colored in light purple, other region of it is colored in cyan. Switch I of chain A of TM EngB/GDP complex partially resembles the conformation of *B. subtilis* EngB /GMPPNP /Mg^{++}. The conformations of the N-terminal end of switch I between the two structures are similar. They are different in the C-terminal half of switch I. Switch II in both structures are not defined indicating the

flexibility of switch II. (b) Residues in switch I region involving interaction with the bound ligands and other region of EngB are shown in sticks and labeled (*B. subtilis* EngB residues are labeled in brackets). Residues involved in interactions found in both TM EngB/GDP complex and *B. subtilis* EngB /GMPPNP/Mg^{++} complex are labeled in italic. Interactions involved are summarized in table 4.2. The N-terminal end of switch I in both chains have similar conformations involving interaction of Ala48 and Ser 51 in TM EngB (Ala50 and Ser53 in that of *B. subtilis*). In *B. subtilis* EngB /GMPPNP/Mg^{++} complex, interactions of residues Lys58 and and Thr59 with the gamma phosphate of the bound nucleotide and magnesium cation define the conformation of the C-terminal end of switch I. Such interactions are not observed in TM EngB/GDP complex as there are no gamma phosphate and magnesium cation. They cause the differences in switch I conformation between a GDP bound TM EngB and GMPPNP bound *B. subtilis*.

Figure 4.7 Structure alignment between apo-enzyme TM EngB and TM EngB /GDP complex



Apo-enzyme TM EngB and TM EngB /GDP complex are aligned together. Switch I of TM EngB/GDP is colored in red, switch II of it is colored in blue, other region of it is colored in grey. Switch I of apo-enzyme TM EngB is colored in orange, other region of it is colored in green. Most of the loop region of switch I and the whole loop region in switch II of apo-TM EngB are not defined, compare with TM EngB in complex with GDP, which the complete loop is defined in switch I and partial loop is defined in switch II. In apo-enzyme form without a bound ligand, loops in both switches are not defined. It indicates that the binding of nucleotide ligand should define the conformation of both switch I and switch II in TM EngB.

Table 4.1 Summary of interactions involved in recognition of GDP molecule in TM EngB

Interactions (atom on TM EngB – atom on GDP)
Sidechain oxygen of Asp143 – Nitrogen 1 on guanine ring
Sidechain oxygen of Asp143 – Nitrogen 2 on guanine ring
Backbone nitrogen of Ser174 -- Oxygen 6 on guanine ring
Backbone oxygen of Ser174 – Nitrogen 7 on guanine ring
Backbone nitrogen on Asn32 – Oxygen on beta phosphate group
Backbone nitrogen on Gly34 – Oxygen on beta phosphate group
Backbone nitrogen on Lys35 – Oxygen on beta phosphate group
Backbone nitrogen on Ser36 – Oxygen on beta phosphate group
Backbone nitrogen on Ser37 – Oxygen on alpha phosphate group
Backbone nitrogen on Ser51 – Oxygen on alpha phosphate group (chain A only)

Table 4.2 Summary of interactions involving switch I and the bound ligands in chain A of TM EngB /GDP complex and *B. subtilis* EngB/ GMPPNP/ Mg^{++} complex

TM /EngB GDP complex	***B. subtilis* EngB/ GMPPNP/ Mg^{++} complex**
Interactions found in both structures	
Asn40 sidechain nitrogen and oxygen – Ala48 backbone oxygen and nitrogen	Asn41 sidechain nitrogen and oxygen – Ala50 backbone oxygen and nitrogen
Ser51 backbone nitrogen – GDP alpha phosphate oxygen	Ser53 backbone nitrogen – GMPPNP gamma phosphate oxygen
Interactions found only in one of the structures	
Asn32 sidechain nitrogen – Lys 52 backbone oxygen	-
-	Lys58 backbone nitrogen – GMPPNP gamma phosphate oxygen
-	Thr59 backbone nitrogen – GMPPNP gamma phosphate oxygen
-	Thr59 sidechain oxygen – magnesium cation

Chapter Five

5 Purified TM EngB is Active for binding guanine nucleotide but inactive for GTPase hydrolysis activity

5.1 Introduction

From the structure solved from the crystal of TM EngB, there was a GDP molecule bound on each protein molecule of TM EngB. There was no addition of exogenous GDP. GDP bound on TM EngB was from *E. coli* cell during expression. After steps of dialysis and chromatographic purification, GDP was not removed from the protein. Therefore, TM EngB must have a high affinity towards ligand GDP. To study the affinity of TM EngB towards its ligand, dissociation constant of TM EngB with its ligand GDP and analogue GTPγS has to be determined. As GTP might degrade to GDP, a GTP analogue, GTPγS will be used instead of GTP in the analysis. Dissociation constant of TM EngB towards its ligand GDP and analogue GTγP will be found by competitive binding experiments using a fluorescent labeled GDP analogue, mant-GDP (2'/3' – O - (N – Methyl – anthraniloyl) – guanosine – 5' - diphosphate), which contains an additional fluorescent group attached on the 2' or 3' oxygen in the ribose of the guanine nucleotide.

5.2 Studying ligand affinity by competitive binding experiment

To find out the dissociation constant of TM EngB towards its ligands, a method to detect the amount of bound ligand and protein is required. To do so, mant-labeled GDP was used to detect the binding of mant-GDP to protein TM EngB. Trytophan on TM EngB could be excited with UV night at around 280nm and give out fluorescence with peak at 350nm. When mant-GDP bound on TM EngB, i.e. the mant-group and the protein were in close proximity (<10Å), the fluorescence light from excited trytophan residues my further excite the mant group on the labeled GDP and give out another fluorescence light at longer wavelength with emission peak at 440nm. Such phenomenon was known as fluorescence resonance energy transfer (FRET), which only happens when two fluorescence giving group were in close proximity, i.e. mant-GDP bound to TM EngB. Therefore the enhanced fluorescence at 440nm could represent the amount of mant-GDP bound TM EngB.

The competitive experiment was done according to Lehoux, I. *et al.*, 2003. The mant-GDP was titrated from a 0.5mM stock solution to samples containing 3mL of 0.5μM of TM EngB in various concentrations of GDP or GTPγS. Fluorescence emitted at 440nm was recorded with excitation at 285nm. Detail setup of the experiment was described as section 2.2.9.

The fluorescence intensity given out at 440nm, Y is a combination of fluorescence of

bound and free mant-GDP according to equation 1:

$$Y = FL \times L + (FEL - FL) \times EL \qquad (1)$$

where EL is the concentration of the TM1466-mant-GDP complex, L is the total concentration of the mant-GDP added, FL is the fluorescence intensity of one micromole of the mant-GDP alone and FEL is the fluorescence intensity of one micromole of the TM1466-mant-GDP complex.

FL was measured in a titration experiment in which the mant-GDP was added to the buffer without TM1466.

EL is the solution for a quadratic equation describing the single site binding of mant-GDP to TM1466. It was derived from

The dissociation constant,

$$K_d = \frac{[E]_{free} \cdot [L]_{free}}{[EL]} \qquad (1)$$

$$K_d = \frac{([E]_{total} - [E]_{bound}) \cdot ([L]_{total} - [L]_{bound})}{[EL]} \qquad (2)$$

[E] – enzyme concentration [L] – Ligand concentration

[EL] – enzyme-ligand complex concentration

Assuming single binding site on the protein,

$[E]_{bound} = [L]_{bound} = [EL]$

$$K_d = \frac{([E]_{total} - [EL]) \cdot ([L]_{total} - [EL])}{[EL]} \qquad (3)$$

Rearrange the terms,

$$EL = \frac{([L]_{total} + [E]_{total} + K_d) - \sqrt{([L]_{total} + [E]_{total} + K_d)^2 - 4 \cdot [L]_{total} \cdot [E]_{total}}}{2} \quad (4)$$

$[L]_{total}$: Total ligand concentration $[E]_{total}$: Total enzyme concentration

EL: Enzyme-ligand complex concentration

In the above equation, L can represent the total mant-GDP concentration, E can represent the total TM1466 concentration and K_d can represent the dissociation constant for mant-GDP of the protein. K_d and FEL were then obtained by fitting the fluorescence data from the titration experiments using equation (1) constraining the fit to L, E, and FL.

By solving the equation, the apparent K_d of mant-GDP at various concentration of competitive ligands GDP or GTPγS could be solved. The apparent K_d values for mant-GDP were plotted against GDP concentration to give a secondary plot to yield the Kd for GDP which was found on the negative of x-intercept of the plot. The calculation involved was described by Segel *et al.*, 1993.

5.3 GDP binds to TM EngB with higher affinity than GTPγS

To find out the K_d of TM EngB towards GDP and GTPγS, apparent K_d of TM EngB towards mant-GDP at different concentration of GDP and GTPγS were solved. Mant-GDP was titrated to samples containing 3mL of 0.5μM of TM EngB in buffer with various

concentrations of GDP or GTPγS, 10mM Tris pH7.5, 0.1M NaCl, 1mM EDTA, 10mM Mg^{2+}. Fluorescence emitted at 440nm was recorded with excitation at 285nm. The detailed setup was described in section 2.2.9. Result was plotted as figure 5.1 and 5.2.

The apparent K_d of TM EngB towards mant-GDP under various concentrations of GDP and GTPγS was found by non-linear regression. K_d was plotted against the corresponding concentration to yield a secondary plot (figure 5.3). From the plot, K_d for the corresponding ligand which act as a competitive inhibitor in each case was estimated as the negative of the x-intercept on each plot. The K_d for GDP was estimated to be 0.48μM and that for GTPγS was 5.9μM. TM EngB showed a higher affinity towards GDP than GTPγS.

5.4 TM EngB showed very low intrinsic GTPase activity

GTPase activity of TM EngB was estimated by measuring the release of inorganic phosphate (Pi) by an end-point assay using the EnzChek phosphate kit (Molecular Probe) according to the manual provided by the manufacturer. During hydrolysis of GTP, inorganic phosphate is released. Phosphate release was measured as the difference of phosphate content before and after overnight incubation. Reaction mixtures containing 1μM of TM EngB and 1800μM GTP and a control setup with 1800μM GTP were prepared

with reaction buffer, 20mM Tris pH7.5, 0.1M NaCl, 1mM EDTA, 10mM Mg^{2+}. 89μM of free phosphate was found in the reaction mixture after 16 hours incubation. <1μM phosphate release was found in control without addition of protein EngB.

5.5 **Discussion**

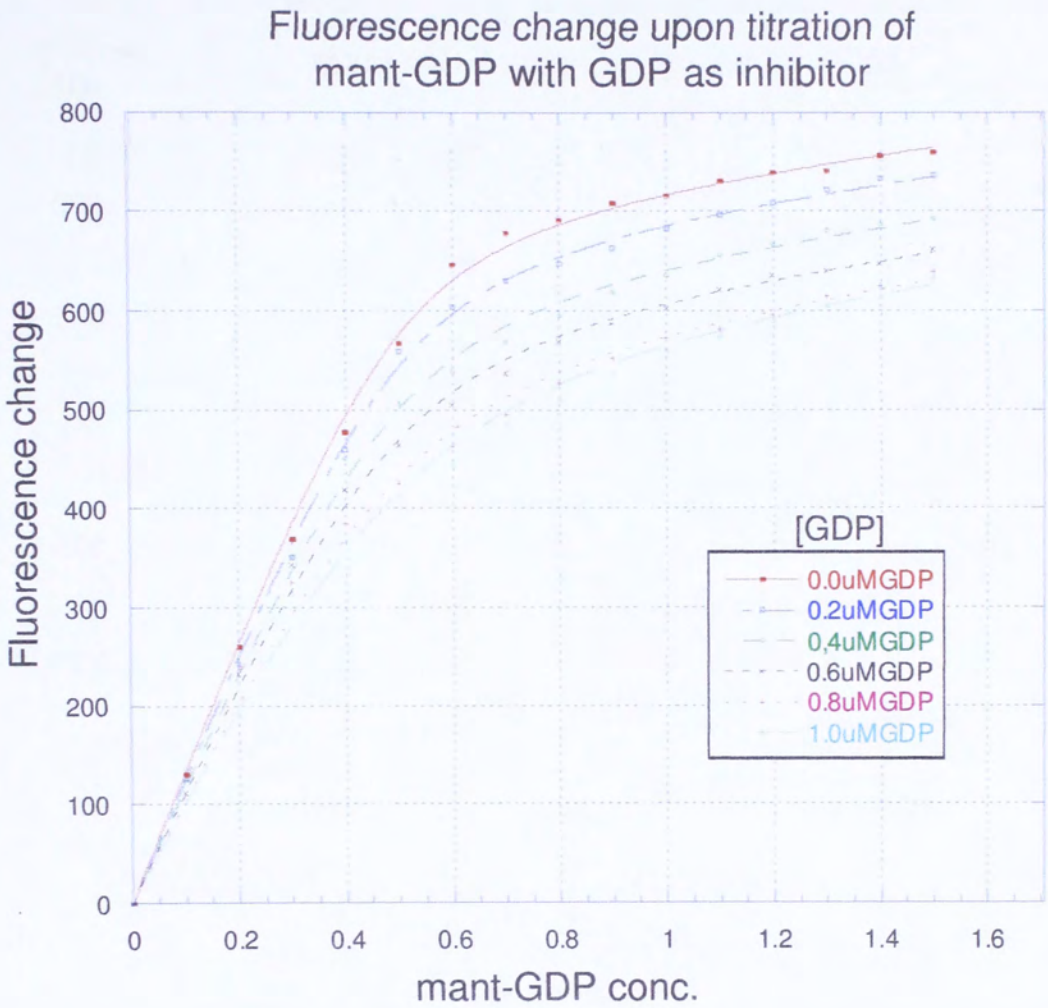
From the affinity experiments, affinity of TM EngB towards its guanine nucleotide ligand was found by competitive titration with fluorescent labeled mant-GDP. The fluorescence given out was enhanced upon adding of TM EngB to the fluorescent labeled ligand analogue. The fluorescence given out was corresponding to the amount of TM EngB / mant-GDP complex. With such information, the apparent dissociation constant (K_d) of TM EngB towards mant-GDP could be found under various concentration of competing guanine nucleotide.

From the experiments, we found that the dissociation of constant of TM EngB towards GDP was found to be 0.48μM and that towards GTPγS, a GTP analogue, was 5.9μM. The affinity towards GDP was higher than that of GTPγS. TM EngB bound diphosphate ligand more readily than triphosphate ligand. Besides, the affinity between TM EngB is high with Kd in submicromolar range, compare with Kd of *E.coli* EngB (3.9μM) [Lehoux, I.E. *et al.*, 2003].

The titration method used to estimate the binding affinity has its own limitation. With high affinity, the determination of K_d is difficult. K_d is found by [Ligand(free)]x[Enzyme(free)] / [Enzyme-ligand complex]. In the setup, only [Enzyme-ligand complex] is found directly by determination of fluorescence given out. At low enzyme-ligand complex concentration, the error comes from the error in detecting weak fluorescence from the complex. At higher enzyme-ligand complex concentration, where [Enzyme-lignad] is comparable to [Ligand(total)] or [Enzyme(total)], error comes from the determination of [Ligand(free)] or [Enzyme(free)] as the Free concentration is determined by the "Total concentration –complex concentration". Though the error in the setup may be great, but the order of magnitude of the affinity constant can still be estimated through this method.

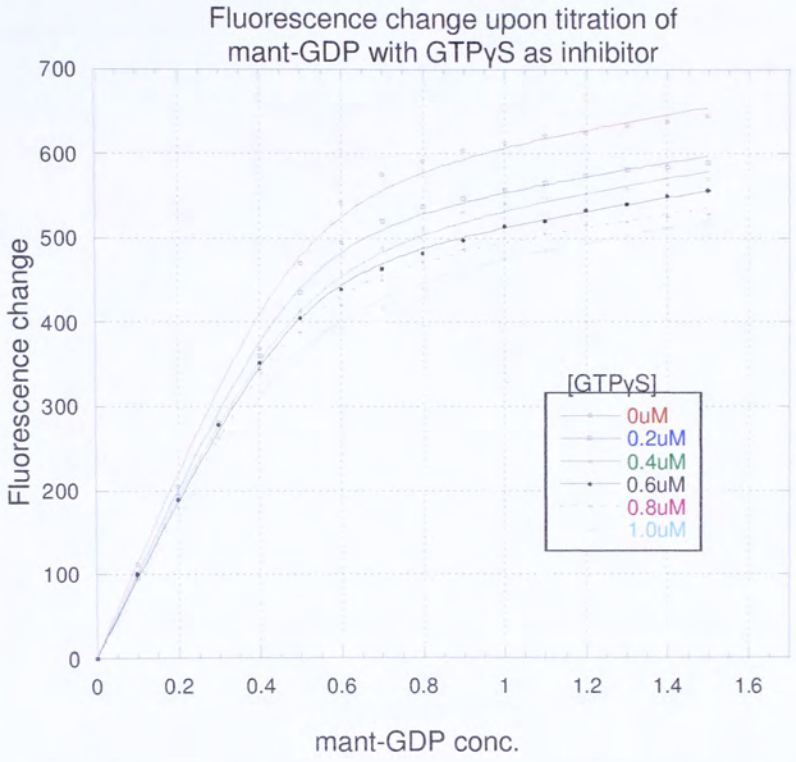
Very low intrinsic hydrolysis activity was detected by incubating TM EngB with GTP. It agrees with the observation in *E. coli* YihA [Lehoux, I.E. *et al.*, 2003]. Such observation may due to lack of other factors allowing hydrolysis of GTP or releasing the bound GDP in TM EngB. G-proteins may require GTPase activating protein to carry out hydrolysis of bound ligand, or guanine nucleotide exchange factors to release the bound GDP. It suggests that other protein may take part for the proper function, switching in a cycle of "on" and "off", of TM EngB.

Figure 5.1 Competitive binding of fluorescent labeled GDP to TM EngB with GDP



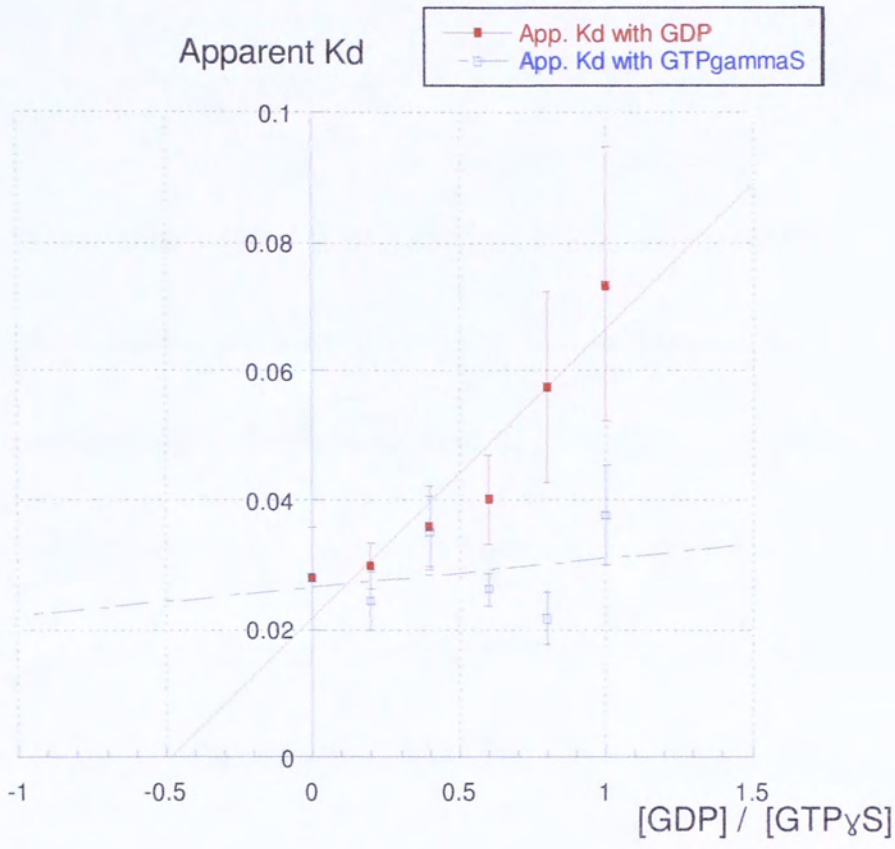
Competition of mant-GDP by GDP. 0.5uM of TM EngB was titrated with mant-GDP with various concentration of GDP. The apparent Kd of mant-GDP at different concentration of GDP was found by fitting the data point by non-linear regression.

Figure 5.2 Competitive binding of fluorescent labeled GDP to TM EngB with GTPγS



Competition of mant-GDP by GTPγS. 0.5uM of TM EngB was titrated with mant-GDP with various concentration of GTPγS. The apparent Kd of mant-GDP at different concentration of GTPγS was found by fitting the data point by non-linear regression.

Figure 5.3 Dissociation constants of GDP and GTPγS are estimated by competitive binding experiments



Competition of mant-GDP with GDP or GTPγS. The apparent Kd of mant-GDP at different concentration of GDP (red line) or GTPγS (blue line) were plotted against their corresponding concentration. The K_d value for each competitive ligand was given by the x-intercept. The K_d for GDP was estimated to be 0.48μM and that for GTPγS was 5.9μM. TM EngB showed a higher affinity towards GDP than GTPγS.

Chapter Six

6 Thermostability of EngB of *T. maritima*

6.1 Introduction

Thermotoga maritima, a rod-shaped bacterium belonging to the order Thermotogales, was originally isolated from geothermal heated marine sediment at Vulcano, Italy. As the organism has an optimum growth temperature of 80°C, its protein should be able to remain stable and functional at such high temperature. To test if protein TM-EngB is thermostable, we performed circular dichroism spectrophotometry experiments to demonstrate thermostability of the protein, in which TM EngB was unfolded by heat or chemical denaturant, and CD signal at 222nm was taken as a measurement of secondary structure. Data was fitted into a two-state model to describe the state of unfolding. Structural comparison of EngB from thermophilic *T. maritima* and mesophilic *E. coli* may provide clues to explain thermostability of *T. maritima* EngB.

6.2 Guanidine hydrochloride – induced unfolding

Conformational stability of TM EngB was studied by chemical unfolding. Guanidine hydrochloride was used as denaturant to unfold TM EngB. 50μM of TM EngB was incubated in 20mM phosphate pH7.5 with 1.4 - 7.2M guanidine hydrochloride (0.1M

intervals from 4.4-5.4M, 0.2M intervals for other concentration) overnight.

Each sample was taken for CD measurement as described in section 2.2.8. CD signal at 222nm was taken as a measurement of presence of secondary structure of protein. The result was fitted into a two-state model as described in section 2.2.8.2 (figure 6.1). [D]½, at which the fraction of folded protein = fraction of unfolded protein = 0.5, was found to be 4.83M. And the free energy of unfolding at zero denaturant concentration at 25°C, i.e. Δ Gu(H2O), was found to be 109kJ/mol.

Reversibility of unfolding TM EngB with GdnHCl was tested by diluting the unfolded TM EngB in 7.2M GdnHCl by ten fold. CD spectrum from 260 to 210 was taken after dilution and compared to the spectrum at 0M, 0.72M and 7.2M GdnHCl (figure 6.2). TM EngB was unfolded at 7.2M GdnHCl. The spectrum after 2 fold dilution resembled the spectrum at 3.6M GdnHCl. It was refolded and regained its secondary structure after 2 fold dilution to 3.6M.

6.3 Thermal-induced unfolding

Conformational stability of TM EngB was studied by thermal unfolding. TM EngB

was unfolded by heating at constant rate. 0.1mM TM EngB in 20mM phosphate buffer was incubated in 60°C for 30 minutes. The protein sample was heated at 60 to 110°C at a rate of 1°C per minute. CD signal at 222nm was recorded. Detail setup was described in chapter 2.2.8.3. As shown in figure 6.3, TM EngB started to unfold at about 88°C. TM EngB was unfolded irreversibly and precipitated out at above mentioned condition. The apparent melting temperature was found to be 91.3°C.

6.4 Structural comparison of thermophilic and mesophilic EngB

Global fold

The two homologues from *T. maritima*, and *E. coli* share similar folds. Major difference is *E. coli* EngB contains an extra C-terminal unstructured peptide (figure 6.4).

Ion pairs

We used the criteria of Szilagyi and Zavodszky (2000) to classify ion pairs using three distance limits of 4 Å, 6 Å, and 8 Å. *T. maritima* EngB contains more salt bridges (ion pairs within 4 Å) than homologue in *E. coli.* Extra salt bridges may account for the thermostability of *T. maritima* EngB. However, sequence comparison between *T. maritima* and *E. coli* EngB revealed that the thermophilic homologue does not contain more charged residues. The *T. maritima* protein has 21 acidic (Glu+Asp) and 32 basic residues (Lys+Arg),

and the *E. coli* homologue has 27 acidic and 27 basic residues. The overall charge of *T. maritima* EngB is -11 to -12 (dependent on the pK_a value of the single histidine residue), and that of *E. coli* homologue is zero to -3 (dependent on the pK_a value of the three histidine residues). Extra salt bridges observed are the result of distribution of charged residues rather than more charged residues presence.

Proline

Proline residues located in loop regions may contribute to protein stability by entropically destabilizing the denatured state [Matthews *et al.*, 1987]. *T. maritima* EngB has 10 proline residues, while *E. coli* homologue has 7. Proline residues observed in *E. coli* homologue are also present in *T. maritima*. By comparing the distribution of proline residues in homologue of *T. maritima* and *E. coli,* extra proline residues of *T. maritima* EngB are in consecutive sequence of two (Pro116 and Pro117) or three (Pro18, Pro 19 and Pro 20) (figure 6.4).

Accessible surface area

Hydrophobic interactions, one of the major driving forces for protein folding, can be correlated with the amount of accessible surface area buried with folding [Makhatadze and Privalov, 1995; Pace, 1995; Janin, 1997]. The solvent accessible surface area was

calculated for the three homologueous proteins by the NACCESS program. The area buried with folding was estimated by subtracting the surface area calculated for the folded state from those for the Ala-X-Ala tripeptide, which serves as a model for the unfolded state [Hubbard *et al.* 1991]. The three homologueous EngB from *T. maritima* , *B. subtilis* and *E .coli* buried a similar amount of total solvent accessible area (~18000 $Å^2$) with folding (Table 6.2). However, the relative amount of polar and nonpolar surface buried are different among the two EngB, with *T. maritima* homologue buries the leaset polar atoms and the most nonpolar atoms with folding (Table 6.2).

Hydrogen bonds

E. coli EngB has slightly less backbone – backbone and backbone- side chain hydrogen bonds than *T. maritima* (table 6.2). More side chain –side chain hydrogen bonds are observed in *E. coli* EngB. Extra side chain – side chain hydrogen bonds are found in the surface residues, Gln 174, Glu176 and Lys 191 (figure 6.5). These residues are replaced by Thr 168 and non-polar Ile 170 and Leu 186 respectively, forming a hydrophobic pocket.

6.5 Discussion

T. maritima EngB is a thermostable protein which starts to denature at 88°C under heat denaturation with an apparent melting temperature of 91.3°C. EngB in *E. coli*, yihA

has a melting temperature of 55°C [Lehoux, *et al.*, 2003], the melting temperature of *T. maritima* EngB is much higher. It is not surprising as *Thermotoga maritima* is a hyperthermophilic bacteria which has a growth temperature at 80°C.

Thermostability of *T. maritima* EngB has been demonstrated. With the structure of several EngB solved (from *T. maritima* and *E. coli*), it is possible for us to analyze thermostability of *T. maritima* EngB through structural study.

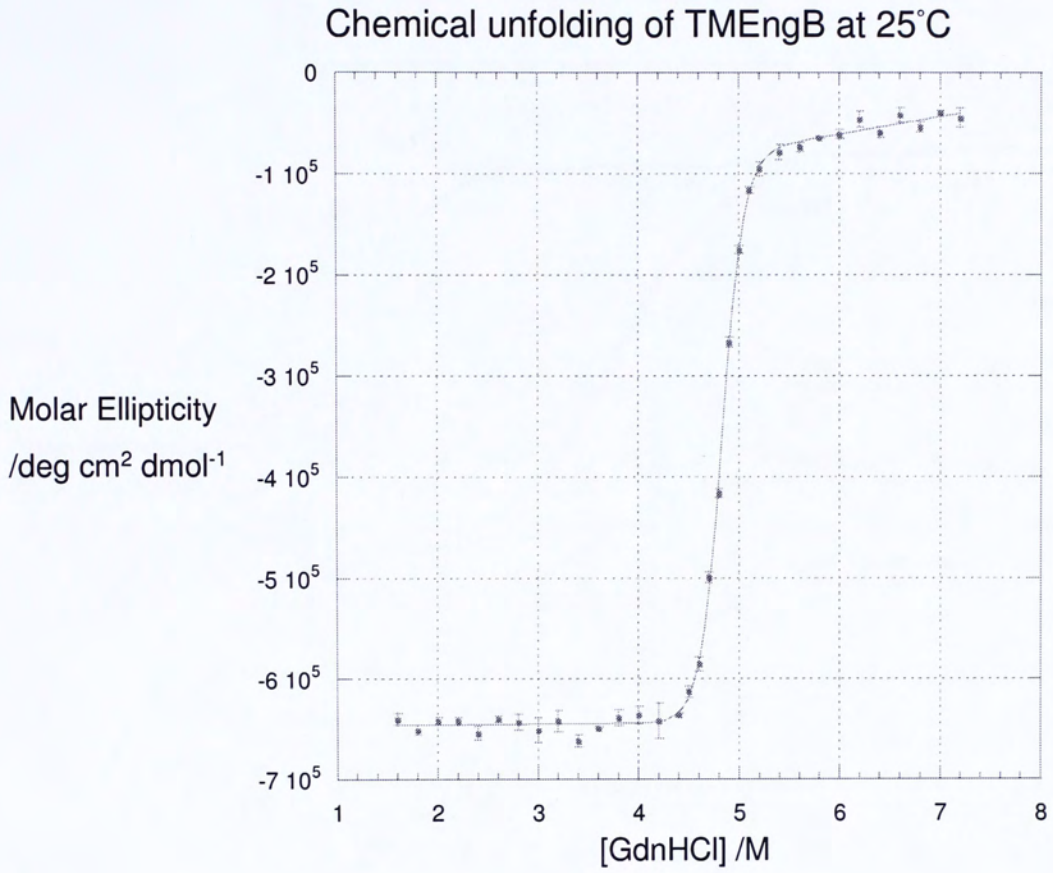
One of the factor observed that may contribute to the thermostability of *T. maritima* EngB is the presence of extra proline residues in loop region of the protein. Proline residues in the loop region may restrict the N-Cα rotations. It has been proposed that proline residues may contribute to protein stability by decreasing the entropy of the unfolded state [Matthews *et al.*, 1987]. The possible conformation of proline is more limited compare with other amino acids, especially glycine. The entropy given up by proline from the unfolded state to folded state of the protein is smaller compared with other amino acids. Hence, the entropy loss during folding of protein is reduced and so folding of protein is more favored. *T. maritima* EngB contains 5 extra proline residues compare with the *E. coli* homologue. Among those extra proline residues, some of them are in consecutive sequence of two (Pro116 and Pro117) or three (Pro18, Pro 19 and Pro 20). It

may be a strategy for enhanced thermostability in *T. maritima* EngB.

Other factors, including ion pairs, hydrogen bonds and buried accessible area show slight differences between the thermophilic and mesophilic homologues. General conclusion could not be drawn without detailed experiment examining each factor in a case by case manner.

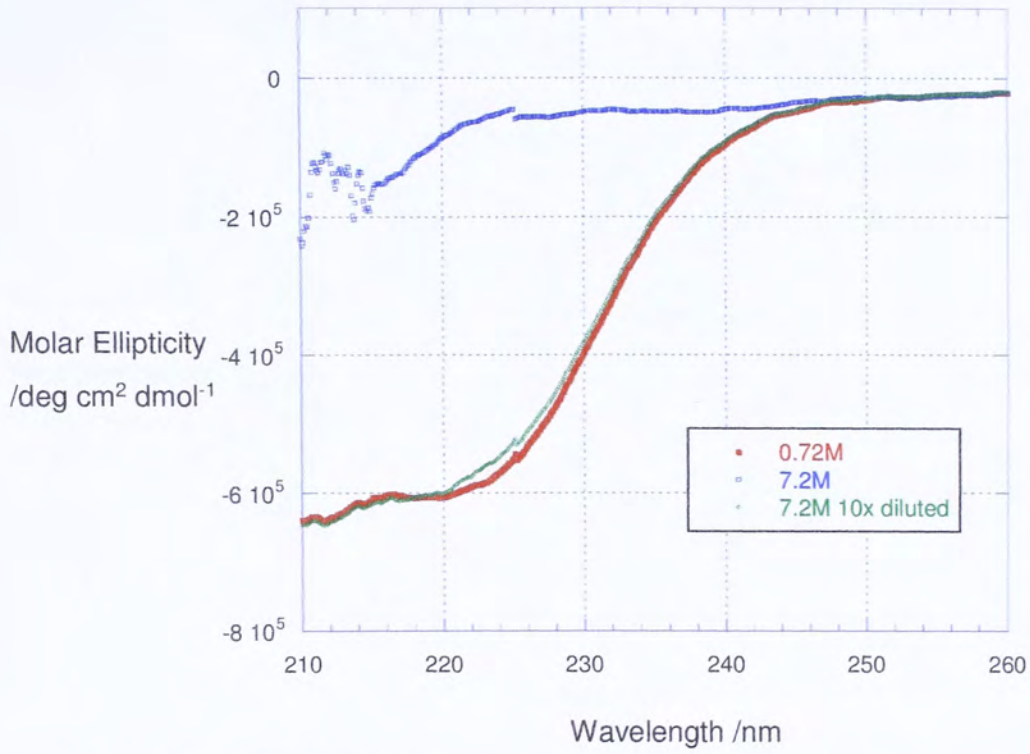
Structural comparison of thermophilic and mesophilic EngB proteins suggests that they do not differ in their packing. Several factors, including presence of extra proline residues, ion pairs, hydrogen bonds and buried accessible surface area may contribute for the thermostability of *T. maritima* EngB. With the protein being capable to be reversibly refolded upon guanidine-induced denaturation, structural features identified in this study can be examined in future experiment. Site directed mutagenesis and thermodynamic measurements can be carried out to analyze the contributions of each feature to thermostability of EngB.

Figure 6.1



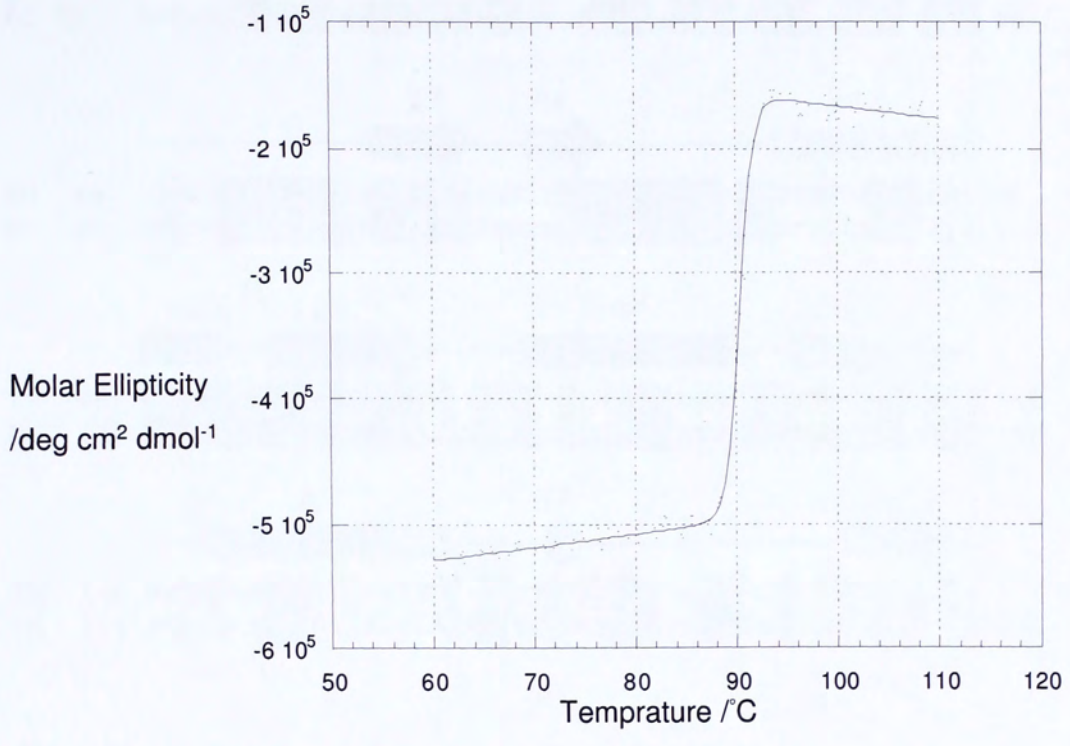
Chemical unfolding of TM EngB, molar ellipticity against concentration of denaturant of GdnHCl. TM EngB started to unfold at 4.4M GdnHCl.

Figure 6.2 CD spectrums at different concentration of denaturant GdnHCl



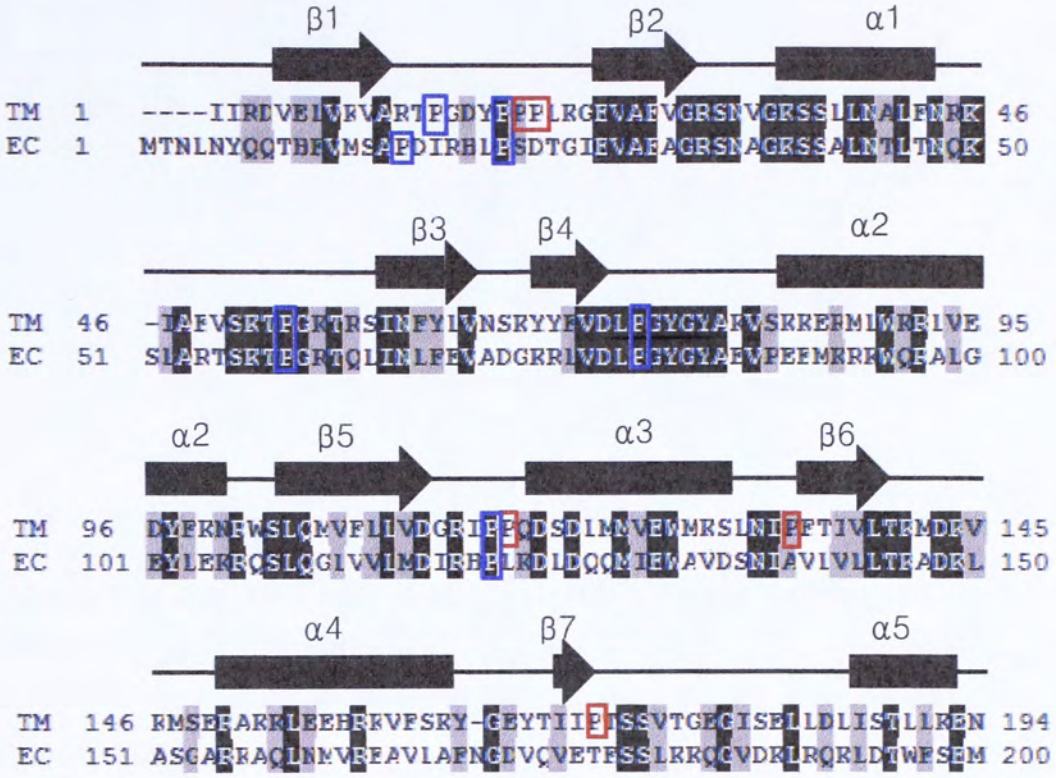
Structural change of TM EngB upon chemical denaturation. Circular dichroism spectrum of TM EngB was obtained from wavelength 210 to 260 at different guanidine hydrochloride concentration. At 0.72M guanidine hydrochloride, TM EngB retained its secondary structure with characteristic negative CD spectrum. At 7.2M guanidine hydrochloride, TM EngB was unfolded with a less negative spectrum. To prove the reversibility of the unfolding by guanidine hydrochloride, a refolding test by simple dilution was done. The sample in 7.2M guanidine hydrochloride was diluted by ten-fold with buffer without denaturant and another spectrum for the diluted sample was measured. The spectrum resembled the spectrum at 0.8M guanidine hydrochloride. The protein regained secondary structure after the ten-fold dilution. The unfolding of TM EngB caused by 7.2M guanidine hydrochloride was a reversible process.

Figure 6.3



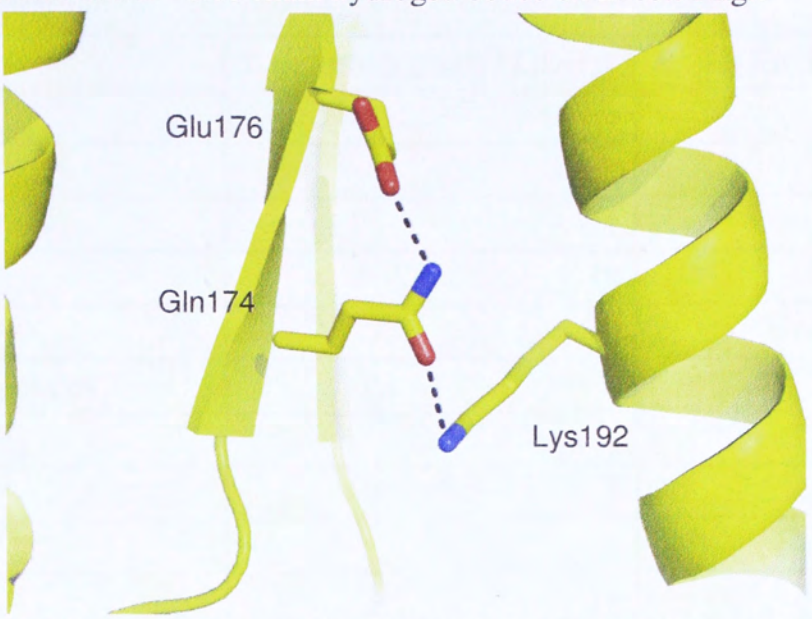
Thermal unfolding of TM EngB. Circular dichroism at 222nm was observed to monitor the change in secondary structure of TM EngB at increasing temperature from 60°C to 110°C(1 °C/min). TM EngB started to unfold at 88°C. TM EngB precipitated at the end of the experiment. It was irreversibly unfolded. The apparent melting temperature (Tm) was found to be 90.3°C.

Figure 6.4 Sequence alignment of EngB from *T. maritima* (TM) and *E. coli* (EC)



Secondary structures are labelled according to TM EngB/ GDP complex structure. Extra proline residues are highlighted with red boxes, conserved proline residues are highlighted in blue boxes. Total of 10 proline residues are present in *T. maritima* EngB and there are 5 in of *E. coli* EngB.

Figure 6.5 Extra side chain – side chain hydrogen bonds on *E. coli* EngB



More side chain – side chain hydrogen bonds are found on *E. coli* EngB. They are found to be on the surface residues Gln 174, Glu176 and Lys 192. Corresponding residues on *T. maritima* are replaced by Thr 168, Ile 170 and Leu 186, forming a hydrophobic pocket at the position.

Table 6.1 Number of ion pairs and hydrogen bonds in *T. maritima*, *B. subtilis* and *E. coli* EngB

	T. maritima EngB / GDP	*E. coli* EngB (apoenzyme)
Ion pairs		
Distance limits		
4Å	11	8
6Å	18	15
8Å	30	29
Repelling charge pairs		
Distance limits		
4Å	1	1
6Å	5	8
8Å	14	17
Hydrogen bonds		
Backbone-Backbone	113	105
Backbone-Side chain	35	26
Side chain-Side chain	15	17
Total	163	140

Number of ion pairs and hydrogen bonds were calculated for the crystal structures of *T. maritima* (chain A of *T. maritima* EngB / GDP complex structure) and compared to *E. coli* EngB (apo *E. coli* EngB [PDB code:1PUI]). Chains of *T. maritima* EngB selected for ion pairs and hydrogen bonds calculation contains most structural information (least missing residues). Undefined missing residues are not involved in the calculation. They are considered flexible in the crystal structures and do not involve in hydrogen bond formation and ion pairs.

Two oppositely charged residues were considered an ion pair if their charged atoms are closer to each other than a certain distance limits. Numbers of ion pairs were counted using three different limits: 4 Å, 6 Å, and 8 Å.

Hydrogen bonds were calculated using the program HBPLUS (McDonald and Thornton 1994). Hydrogen bonds were classified into three groups: backbone-to-backbone, backbone-to-side chain, and side chain-to-side chain.

Table 6.2 Comparison of burial of accessible area

	T. maritima EngB	*E. coli* EngB
All-atomsΔASA	18870Å^2	18208 Å^2
Non-polar ΔASA	12036 Å^2 (64%)	10957 Å^2 (60%)
Polar ΔASA	6834 Å^2 (36%)	7252 Å^2 (40%)
Total no. of residues in calculation	168	169

Solvent accessible surface areas were calculated by the NACCESS program [Hubbard and Thornton 1993] for the crystal structures of *T. maritima*, *B. subtilis* and *E. coli* EngB (chain A of *T. maritima* EngB / GDP complex structure and apo *E. coli* EngB [PDB code:1PUI]). Chain A of *T. maritima* EngB is selected for ASA calculation which contains most structural information (least missing residues). To simplify the comparison, all bound ligands and the residues in *T. maritima* EngB correspond to the unstructured flexible loop in *E. coli* EngB crystal structure are removed before ASA calculation (therefore only the core structured regions of the proteins are compared). A probe radius of 1.4 Å was used in all calculations. Buried surface areas (△ASA) were obtained by subtracting the solvent accessible area calculated for the folded state from those for the tripeptide Ala-X-Ala, which serves as a model for the unfolded polypeptide [Hubbard *et al.* 1991]. Accessible surface areas for atoms N and O are considered polar and all other atoms are nonpolar. Missing residues in crystal structures are not involved in ASA calculation. Number of residues observed and missing in each peptide chain is listed as reference.

Chapter Seven

7 Construction of a dual-tag affinity pull-down system for finding interacting partner of EngB

7.1 Introduction

Biochemical study of family member, EngB in *E. coli*, suggested that this group of protein can actually bind to guanine nucleotide [Lehoux et al., 2003]. Morevoer, EngB in *E. coli* and *B. subtilis* were found to be essential in corresponding species in knock-out experiments [Arigoni et al., 1998; Dassain et al., 1999; Pragai and Harwood, 2000]. In the knock-out experiments of EngB in *E. coli*, cell cycle defect was observed [Dassain et al., 1999]. Therefore, it is speculated that the EngB family is related to cell cycle process. However, no direct prove had been discovered showing the actual function of EngB. Analyzing the interacting proteins of EngB can help us to find out the function of EngB. Therefore, we have to isolate those interacting proteins.

To demonstrate the feasibily of isolating interacting proteins with EngB by *in vitro* pull-down, a dual-tag pull down system is developed. *E. coli* was chosen as the study model in pull-down assay. A dual-tag system was developed which overexpress tagged EngB to acquire a large amount of bait protein. With abundant tagged EngB as bait, chance

to fish out interacting protein was higher. With a dual tagged EngB, the background of non-specific binding would be reduced.

The dual-tag system is developed with basis on the tandem affinity purification (TAP) system from Strategene. The dual tags used are streptavidin binding peptide (SBP) and calmodulin binding peptide (CBP). Both tags have the benefit of high specificity and high affinity. Binding and elution of both tags are also mild which does not destroy the interaction between our bait EngB with the interacting partners.

7.2 Preparation of dual-tagged *E.coli* EngB

7.2.1 Cloning of SBP-CBP-EC EngB expression construct

The expression construct was pRSET based. The SBP-CBP encoding DNA fragment was acquired by PCR with SBP-CBP specific primers from the pNTAP-1 plasmid from Strategene. A thrombin recognition site (LVPRGS) fragment was engineered into the C-terminal of the fragment adding the sequence to the primer. The fragment was cloned into multiple cloning site on pRSETA plasmid vector utilizing restriction site *Nde*I and *Bam*HI. Cloning procedures was performed as described in section 2.2.2. The resulted plasmid was named pRSET-SBP-CBP.

E.coli EngB encoding DNA fragment was acquired by PCR with *E.coli* EngB specific primers from the *E.coli* (DH5α) lysate. The fragment was cloned into multiple cloning site

on pRSET-SBP-CBP plasmid vector utilizing restriction sites *Bam*HI and *Eco*RI. Cloning procedures was performed as described in section 2.2.2. The resulted plasmid was named pRSET-SBP-CBP-EC EngB.

7.2.2 Expression and purification of SBP-CBP-EC EngB

Plasmid pRSET-SBP-CBP-EC EngB was transformed into *E. coli* BL21(DE3)pLysS. Fusion SBP-CBP-EC EngB fusion protein was produced by expressing the protein in *E. coli* as described in (section 2.2.3). The cells were harvested and lysed. Proteins in the cells were analyzed by SDS-PAGE. SBP-CBP-EC EngB (32kDA) was found to be over-expressed with a major band shown in SDS-PAGE analysis (figure7.1).

SBP-CBP-EC EngB was purified by SP-ion exchange chromatography and Heparin affinity chromatography was finally polished by running gel filtration chromatography as described in section 2.2.4 (figure 7.2).

7.3 Pull down using dual tagged *E.coli* EngB as bait to isolate potential interacting partners of EngB

Pull down experiment was performed to isolate the potential interacting partners of *E.coli* EngB. SBP-CBP- tagged EC EngB was loaded on to streptavidin resin. Protein from

E.coli lysate was allowed to bind on the EngB bound streptavidin resin. Two control experiments was set up, one without EngB and another without *E.coli* lysate protein. Unbound protein was washed away. The bound proteins were eluted and loaded on to calmodulin resin. Unbound protein was washed and the bound protein was eluted and analysed by SDS-PAGE. Detail experimental setup was described in section 2.2.10.

The flow-through and washing fractions were analyzed by SDS-PAGE and stained by Coomassie Blue (figure 7.3). The elution fractions were analyzed by SDS-PAGE and stained by silver stain. From the pull-down experiment, potential interacting partners of *E.coli* EngB could be isolated.

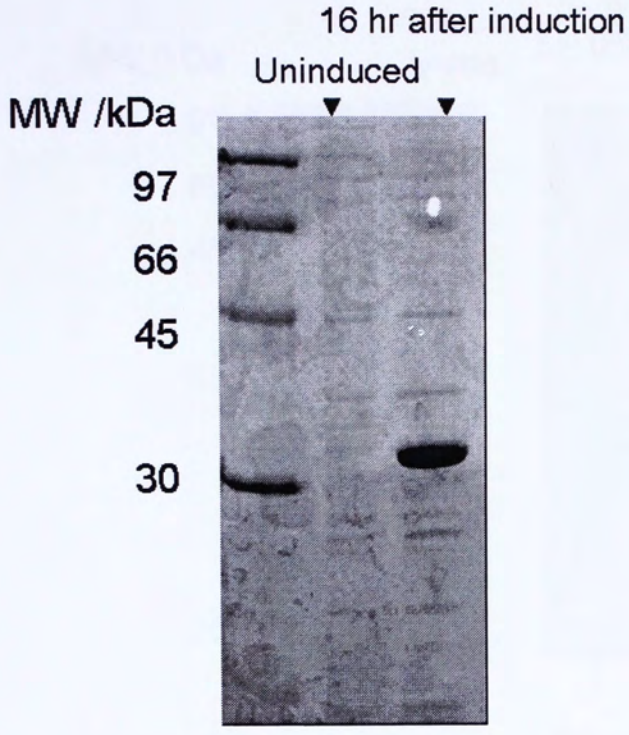
7.4 Discussion

A dual – tag pull down system has been established to study interacting proteins. Feasibility of isolating potential interacting partners of EngB by *in vitro* pull-down was demonstrated. To complete the experiments, another negative control using the SBP-CBP tag as bait will be included and those isolated proteins will be identified.

In parallel to our work, *Bacillus subtilis* EngB had been demonstrated to be associated with the large ribosomal subunit by ultracentrifugation [Schaefer, L. *et al.*, 2006]. It was expected that those proteins co-purified with EngB as a protein complex after the 2-step

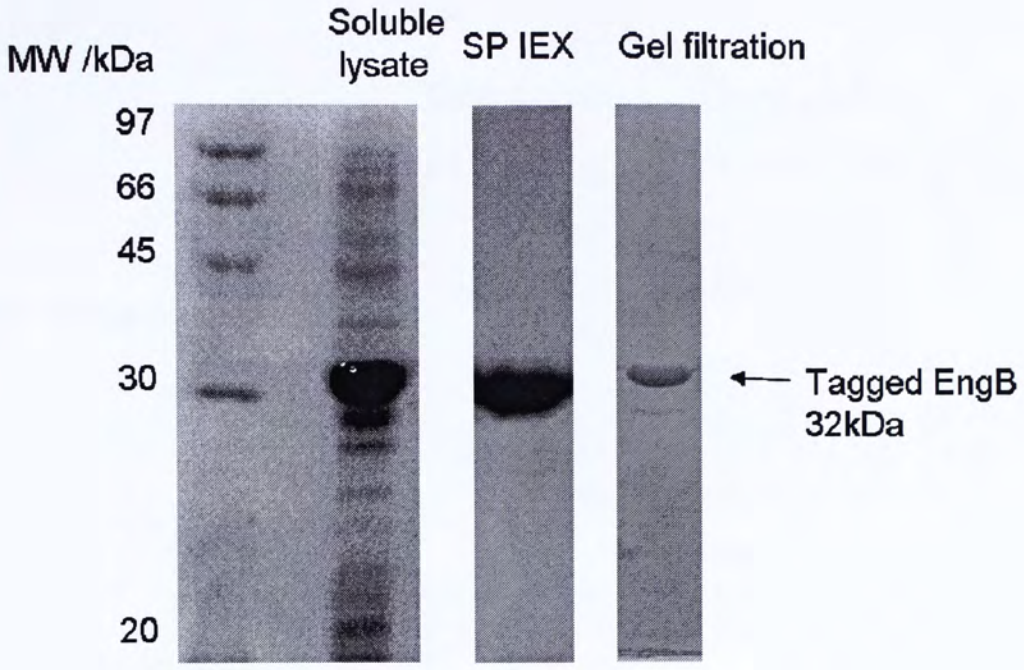
pull-down could be ribosome related proteins. However, actual function of EngB is still not clear. To identify those proteins, mass spectroscopy may be applied. With the identity of the pulled-down proteins, it is possible to predict the function of EngB. Besides, by obtaining co-crystal structure of EngB with its interacting proteins, we can understand the structural mechanism of interactions of EngB with other proteins. It gives us information of specific drug design which disrupts the function of EngB for new antibiotics.

Figure 7.1



E. coli EngB with N-terminal dual-tag of streptavidin binding peptide (SBP) and calmodulin binding peptide(CBP) was expressed in soluble fraction using *E.coli* BL21 (DE3) pLysS. Protein profiles of the expression cells before and after induction of expression were analyzed by SDS-PAGE and stained by Comassie blue. After the expression culture reached OD600=0.4, IPTG was added to 0.4mM concentration to induce expression of the fusion protein. 16 hours after induction, cells were collected by centrifugation and were disrupted by sonication.

Figure 7.2



E. coli EngB with N-terminal dual-tag of streptavidin binding peptide (SBP) and calmodulin binding peptide (CBP) was expressed in soluble fraction using *E.coli* BL21 (DE3) pLysS. Protein in each phase of purification was analyzed by SDS-PAGE and stained by Coomassie blue. Cells were collected by centrifugation and were disrupted by sonication. Lysate was centrifuged and the supernatant which contains the soluble fraction of protein was collected for further purification. Fusion SBP-CBP EngB (32kD)was purified by SP ion exchange column. The tagged EngB was further purified by gel filtration column.

Figure 7.3a Binding, washing and elution profiles of affinity pull-down experiment of *E.coli* EngB

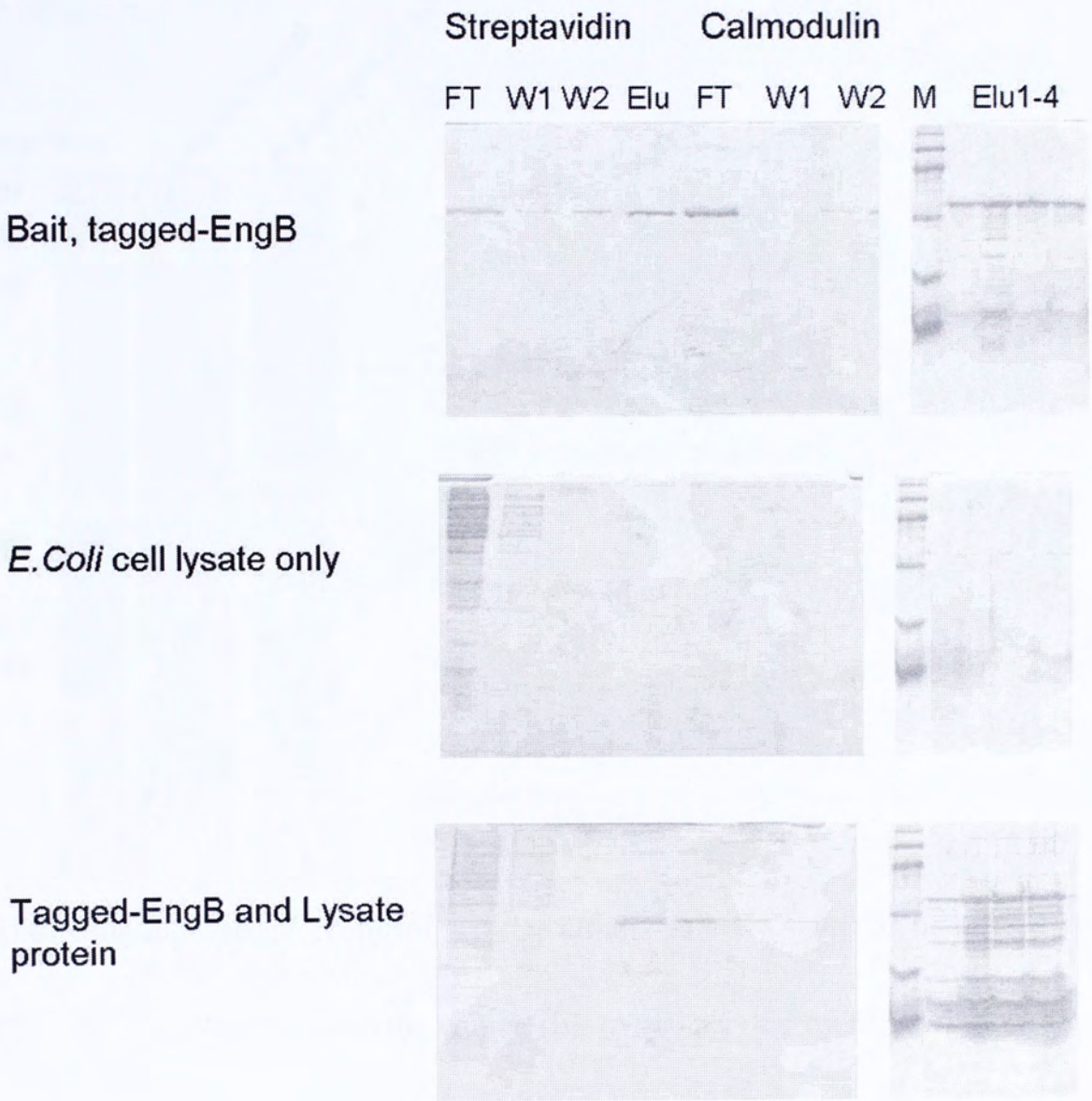
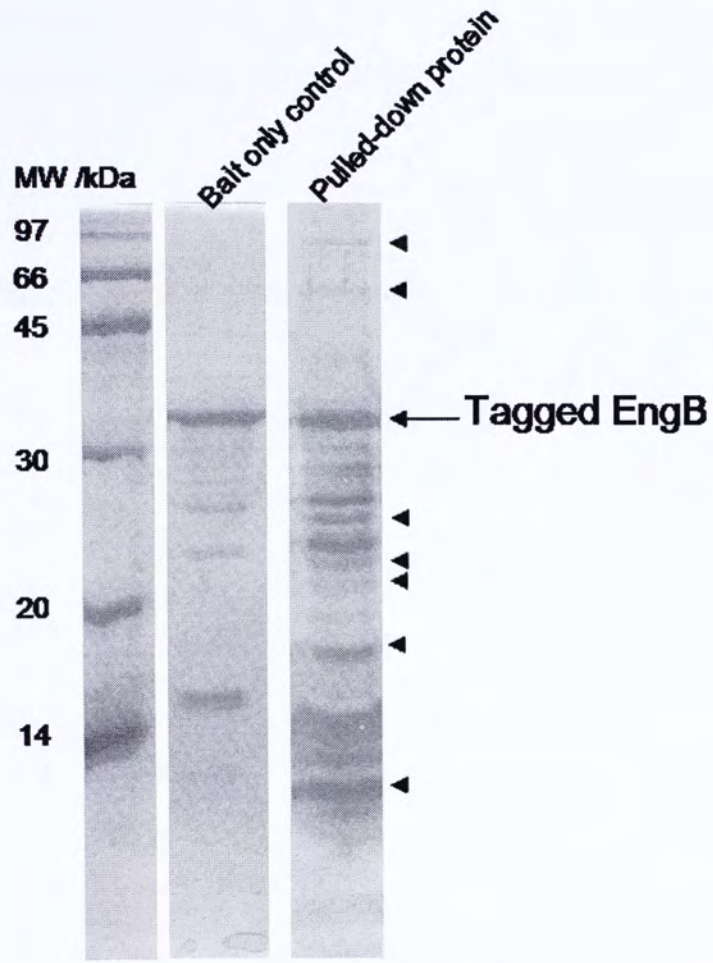


Fig 7.3b



(a) Binding and washing profile of two-step affinity pull-down. Protein in the flow through (FT), washing (W1, W2) and the elution (Elu) was analyzed by SDS-PAGE and stained with Commassie blue. It showed that the tagged EngB was capable to bind on the streptavidin and calmodulin resin. The elution fractions in the pull –down assay were analyzed by SDS-PAGE stained with silver stain, total of 4 elution fractions were collected for each setup. No proteins were detected in the lysate only control without tagged EngB added. (b) Elution profile of pull-down with tagged-EngB was compared with the bait-only control. Proteins found only in the pull-down experiment were marked with arrow. Those

proteins were remaining bound on EngB after pull-down utilizing the SBP and CBP tag. Several other proteins were observed in the elution fraction of both setup. Those proteins were co-purified with EngB and were still present after 2-steps pull-down. It was speculated that they formed complex with the tagged-EngB protein since EngB was translated and so they were also potential interacting partners of EngB.

Chapter Eight

8 Conclusion

From our study of EngB, several conclusions have been drawn. First, the switch I conformation of EngB/GDP complex is found to be not completely disordered as previously reported in *B. subtilis* EngB/GDP complex. Crystal structure of *T. maritima* EngB is solved in apo form and GDP-bound form in this work. In the crystal structure of TM EngB/GDP complex, the full switch I loop is detected in one of the two independent monomers observed in asymmetric unit; it is not detected in another monomer, which is like the observation in *B. subtilis* EngB/ GDP complex. One of the possible conformations is observed in the structured switch I loop chain, while the possibility of more than one conformation in the region is proven by another undefined switch I loop chain. The defined switch I loop resembles the conformation observed in *B. subtilis* EngB/ GMPPNP /Mg^{++} complex. The conformation of switch I loop observed in GDP-bound state has never been reported in other EngB structures.

Second, EngB is a monomer in solution in apo form, GDP-bound state and GMPPNP-bound state. It is supported by light-scattering data which measures the native molecular weight of EngB in solution.

Besides, TM EngB has a high affinity towards guanine nucleotide. The dissociation

constant of TM EngB towards guanine nucleotide was measured to be in micromolar to submicromolar range by competitive binding titration. It is not likely that TM EngB may exchange its bound nucleotide on its own. Low intrinsic GTP hydrolysis activity suggests this protein may require other factors for its activity. GAPs (GTPase activating protein) and GEFs (guanine nucleotide exchange factors) may take part in the hydrolysis of GTP and exchange of bound nucleotides for EngB. To understand the function of this protein, finding its interaction partners, including its GAPs and GEFs, can be one of the possible solutions. As a starting step to solve this question, a dual-tag pull-down system was developed to isolate potential interacting partners of *E. coli* EngB. This provides basis to identification of interacting proteins of EngB and reveals its function. Identified interacting proteins can be co-crystallized with EngB to allow study of its interaction mechanism.

From our study, we find that TM EngB is a thermostable protein with an apparent melting temperature of 91°C. Comparing the crystal structure of EngB from *T. maritima* with that from *B. subtilis* and *E. coli* suggests several structural characteristics that may contribute to the thermostability of TM EngB. Those characteristics can serve as information for designing mutant protein to study thermostability of protein EngB.

References

1. Arigoni,F., Talabot,F., Peitsch,M., Edgerton,M.D., Meldrum,E., Allet,E., Fish,R., Jamotte,T., Curchod,M.L. and Loferer,H. (1998) A genome-based approach for the identification of essential bacterial genes. Nat.Biotechnol., 16, 851-856.

2. Blessing, R. H. (1995). An empirical correction for absorption anisotropy. Acta Cryst. A51, 33-38

3. Bourne H.R., Sanders D.A., McCormick F. (1991) The GTPase superfamily: conserved structure and molecular mechanism. Nature 349(6305):117-27.

4. Brunger, A. T., Adams, Paul D., Clore G. M., DeLano W. L., Gros P., Grosse-Kunstleve, R.W., Jiang J., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., and Warren,G.L. (1998). Crystallography & NMR System , Acta Cryst. D54, 905-921

5. Caldon, C. E., Yoong, P. and March, P. E. (2001) Evolution of a molecular switch: universal bacterial GTPases regulate ribosome function. Molecular Microbiology 41(2), 289-97.

6. Caldon, C.E. and March P. E. (2003) Function of the universally conserved bacterial GTPases. Curr Opin Microbiol. 6(2):135-9.

7. COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. (1994). The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst. D50, 760-763

8. Dassain,M., Leroy,A., Colosetti,L., Carole,S. and Bouche,J.P. (1999) A new essential gene of the 'minimal genome' affecting cell division. Biochimie, 81, 889-895.

9. Diederichs K. and Karplus P. A.(1997), Improved R-factors for diffraction data analysis in macromolecular crystallography, Nature Structural Biology, 4, 269-275

10. Emsley, P. and Cowtan, K. (2004). Coot:Model-Building Tools for Molecular Graphics, Acta Cryst D69, 2126-2132

11. Evans P.R., "Data reduction", Proceedings of CCP4 Study Weekend, (1993), on Data Collection & Processing, pages 114-122

12. Evans P.R., "Scaling of MAD Data", Proceedings of CCP4 Study Weekend, (1997), on Recent Advances in Phasing

13. Fraser, C.M., Gocayne, J.D., White, O., Adams, M.D., Clayton, R.A., Fleischmann, R.D., *et al.* (1995) The minimal gene complement of Mycoplasma genitalium. Science 270: 397-403.

14. French, S. & Wilson, K. (1978). On the treatment of negative. intensity observations. Acta Cryst. A34, 517-525

15. Hamilton, W.C., Rollett, J.S. and Sparks, R.A. , Acta Cryst. (1965) On the relative scaling of X-ray photographs, 18, 129-130

16. Hooft R.W.W., Vriend G., Sander, C. and Abola, E.E. (1996). Errors in protein structures Nature 381, 272

17. Hubbard, S.J., Campbell, S.F., and Thornton, J.M. (1991) Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. J. Mol. Biol. 220: 507-530.

18. Janin, J. (1997) Angstroms and calories. Structure 5: 473-479.

19. Kabsch W. (1988) Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. J.Appl.Cryst. 21, 916-924

20. Lehoux,I.E., Mazzulla,M.J., Baker,A. and Petit,C.M. (2003) Purification and characterization of YihA, an essential GTP-binding protein from Escherichia coli. Protein Expr.Purif., 30, 203-209.

21. Leipe DD, Wolf YI, Koonin EV, Aravind L. (2002) Classification and evolution of P-loop GTPases and related ATPases. J Mol Biol. 15;317(1):41-72.

22. Makhatadze, G.I. and Privalov, P.L. (1995) Energetics of protein structure. Adv. Protein Chem. 47: 307-425.

23. Manfred W., (2001) Global Indicators of X-ray data quality, J.Appl.Cryst. 34, 130-135

24. Matthews, B.W. (1968). Solvent content of Protein Crystals, J. Mol. Biol. 33, 91-497

25. Matthews, B.W., Nicholson, H., and Becktel, W.J. (1987) Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. Proc. Natl. Acad. Sci. 84: 6663-6667.

26. McRee, D. E. (1999). XtalView/Xfit¡Xa versatile program for manipulating atomic

coordinates and electron density. J. Struct. Biol. 125:156-165.

27. Murshudov G., Vagin A. and Dodson E. (1996). Application of Maximum Likelihood Refinement, Refinement of Protein structures. Proceedings of Daresbury Study Weekend.

28. Murshudov G.N., Vagin A.A. and Dodson E.J.(1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. Acta Cryst. D53, 240-255.

29. Murshudov, G.N. , Lebedev , A., Vagin , A.A., Wilson , K.S. and Dodson , E.J. (1999). Efficient anisotropic refinement of Macromolecular structures using FFT, Acta Cryst. section D55, 247-255.

30. Neves S.R., Ram P.T., Iyengar R. (2002). "G Protein Pathways.". Science 296: 1636-9.

31. Nozaki, Y., (1972) The preparation of guanidine hydrochloride. Methods Enzymol. 26, 43-50

32. Pace, C.N. (1995) Evaluating contribution of hydrogen bonding and hydrophobic bonding to protein folding. Methods Enzymol. 259: 538¡V554.

33. Pannu ,N.J., Murshudov ,G.N., Dodson, E.J. and Read, R.J. (1998). Incorporation of Prior Phase Information Strengthen Maximum-Likelihood Structure Refinemen. Acta Cryst. section D54, 1285-1294.

34. Perl D, Schmid FX. (2001) Electrostatic stabilization of a thermophilic cold shock

protein. J Mol Biol. 313(2):343-57)

35. Potterton, E., Briggs, P., Turkenburg, M. and Dodson, E. (2003) A graphical user interface to the CCP4 program suite ,Acta. Cryst. D59 1131-1137

36. Pragai,Z. and Harwood,C.R. (2000) YsxC, a putative GTP-binding protein essential for growth of Bacillus subtilis 168. J.Bacteriol., 182, 6819-6823.

37. Puig,O., Caspary,F., Rigaut,G., Rutz,B., Bouveret,E., Bragado-Nilsson,E., Wilm,M. and Seraphin,B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods, 24, 218-229.

38. Read R. (1999), "Outlier rejection", Proceedings of CCP4 Study Weekend, , on Data Collection & Processing

39. Ruzheinikov,S.N., Das,S.K., Sedelnikova,S.E., Baker,P.J., Artymiuk,P.J., Garcia-Lara,J., Foster,S.J. and Rice,D.W. (2004) Analysis of the open and closed conformations of the GTP-binding protein YsxC from Bacillus subtilis. J.Mol.Biol., 339, 265-278.

40. Schaefer, L, Uicker, W.C., Wicker-Planquart, C., Foucher, A.E., Jault, J.M., Britton R.A. (2006) Multiple GTPases participate in the assembly of the large ribosomal subunit in Bacillus subtilis. J. Bacteriol. 188(23):8252-8.

41. Segel, I.H., Enzyme kinetics, Wiley Classics Library Edition, Wiley, New York, 1993, p. 108-109.

42. Szilagyi, A. and Zavodszky, P. (2000) Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: Results of a comprehensive survey. Structuree 8, 493-504

43. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Research, 11;22(22):4673-80

44. Voet, D.and Voet, J.G.. (2004) Biochemistry 3rd Ed.. John Wilely & Sons, New York.

45. Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program, J. Mol. Graph. 8, 52-56

46. Weiss, M. and Hilgenfeld, R. (1997), On the use of the merging R factor as a quality indicator for X-ray data", J.Appl.Cryst. 30, 203-205

47. Winn , M., Isupov , M. and Murshudov ,G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement, Acta Cryst. D57, 122-133